

SEARCH REQUEST FORM

Access DB#

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____

Art Unit: _____ Phone Number 30 _____ Serial Number: _____

Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

For Sequence Searches Only Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

STAFF USE ONLY		Type of Search	Vendors and cost where applicable
Searcher:	<u>Debra - Daly</u>	NA Sequence (#)	STN _____
Searcher Phone #:	<u>308-4501</u>	AA Sequence (#)	Dialog _____
Searcher Location:	<u>Biotech Lib.</u>	Structure (#)	Questel/Orbit _____
Date Searcher Picked Up:	<u>7/3/01</u>	Bibliographic	Dr. Link _____
Date Completed:	<u>7/3/01</u>	Litigation	Lexis/Nexis _____
Searcher Prep & Review Time:		Fulltext	Sequence Systems <u>ASS502</u>
Clerical Prep Time:	<u>3 mi</u>	Patent Family	WWW/Internet _____
Online Time:	<u>2 mi</u>	Other	Other (specify) _____

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

On protein - protein search, using Bw model
Run on: July 3, 2001, 16:13:43 ; Search time 23 Seconds
Sequence: (without alignments)
Scoring table: 1241.474 Million cell updates/sec

Title: US-09-454-223-2
Perfect score: 2487

Sequence: 1 MPPFLSMLVLLVQPLGNLCA. TDPSSQVSHGTGFTSFGGLKL 471.

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676
Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Post-processing: Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_0501:*

1: /SIDS8/gcdata/geneseq/geneseq/AA1986.DAT:/*
2: /SIDS8/gcdata/geneseq/geneseq/AA1987.DAT:/*
3: /SIDS8/gcdata/geneseq/geneseq/AA1988.DAT:/*
4: /SIDS8/gcdata/geneseq/geneseq/AA1989.DAT:/*
5: /SIDS8/gcdata/geneseq/geneseq/AA1990.DAT:/*
6: /SIDS8/gcdata/geneseq/geneseq/AA1991.DAT:/*
7: /SIDS8/gcdata/geneseq/geneseq/AA1992.DAT:/*
8: /SIDS8/gcdata/geneseq/geneseq/AA1993.DAT:/*
9: /SIDS8/gcdata/geneseq/geneseq/AA1994.DAT:/*
10: /SIDS8/gcdata/geneseq/geneseq/AA1995.DAT:/*
11: /SIDS8/gcdata/geneseq/geneseq/AA1996.DAT:/*
12: /SIDS8/gcdata/geneseq/geneseq/AA1997.DAT:/*
13: /SIDS8/gcdata/geneseq/geneseq/AA1998.DAT:/*
14: /SIDS8/gcdata/geneseq/geneseq/AA1999.DAT:/*
15: /SIDS8/gcdata/geneseq/geneseq/AA2000.DAT:/*
16: /SIDS8/gcdata/geneseq/geneseq/AA2001.DAT:/*
17: /SIDS8/gcdata/geneseq/geneseq/AA2002.DAT:/*
18: /SIDS8/gcdata/geneseq/geneseq/AA2003.DAT:/*
19: /SIDS8/gcdata/geneseq/geneseq/AA2004.DAT:/*
20: /SIDS8/gcdata/geneseq/geneseq/AA2005.DAT:/*
21: /SIDS8/gcdata/geneseq/geneseq/AA2006.DAT:/*
22: /SIDS8/gcdata/geneseq/geneseq/AA2007.DAT:/*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

ALIGNMENTS

RESULT	1
ID	AAR36701 standard; Protein: 261 AA.
AC	AAR36701;
XX	
XX	
XX	
DT	12-AUG-1993 (first entry)
XX	
DE	CD40-L.
XX	
KW	Human; CD40-L; CD40; type II; membrane; polypeptide; extracellular; transmembrane; region; intracellular; soluble; activity; B cell; proliferation; induction; antibody; secretion; IgE; agonist; antagonist; binding assay.
KW	
XX	
OS	Homo sapiens.
XX	
FH	Location/Qualifiers
FT	47..261 /note= "Extracellular region"
FT	
XX	
PN	W09308207-A.
XX	
PD	29-APR-1993.
XX	
PF	23-OCT-1992; 92WO-US08990.
XX	
PR	25-OCT-1991; 91US-0703707.
PR	05-DEC-1991; 91US-0805723.
XX	
PA	(IMMV) IMMUNEX CORP.
XX	
PI	Armitage RJ, Fanslow WC, Spriggs MK;
XX	

DR* WPI: 1993-12417/18.

DR* N-PSDB; AAR39506.

XX PT New cytokine CD40-L as CD40 agonist and antagonist - is used for
PT treating allergies, lupus, rheumatoid arthritis, graft-versus-host disease and insulin-dependent diabetes mellitusXX PS claim 1; Fig 2; 80pp; English.
CC this sequence represents human CD40-L polypeptide which binds to CD40.
CC CD40-L is a type II membrane polypeptide which has an extracellular
CC region at its C-terminus, a transmembrane region and an intracellular region at its N-terminus. A soluble form of CD40-L
CC lacks the transmembrane domain. CD40-L activity is mediated by
CC binding with CD40 an induces B cell proliferation and induction of
CC antibody secretion, including IgE. Membrane bound CD40-L acts as a
CC CD40 agonist and soluble CD40-L acts as a CD40 antagonist. CD40-L
CC can be used in a binding assay to detect cells expressing CD40.

XX SQ Sequence 261 AA;

Query Match 44.6%; Score 1110; DB 14; Length 261;
Best Local Similarity 100.0%; Pred. No. 1.3e-73; Indels 0; Gaps 0;
Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;QY 257 HRRDKIEDRNHLRDFVFMKTIQRCNGERSLSLNEEIKSQFEGFVKDMLNNEETK 316
Db 47 hrrdkiedrnhlrdedfvfmktiqrcngerslslnceekksqfegfvkdlmkeetk 106
QY 317 KENFEMOKGDQDQPIQIAHVVISASSKTSVQIWAEGQYTMNNLVTLENGKQLTVKQ 376
Db 107 kensfemqkdgqpqiahvvisassktsvqiwaeqytmnnlvtlengqqltvkq 166
QY 377 GLYYIAYQVFCNREASSQAPIASLCKSICRERFLILRAANTHSSAKPCGQOIHG 436
Db 167 glyiyiayqvtfcnreassqapiaslickspqrerflilaanttssakpcggqshig 226
QY 437 GVFELQPGASVFNWNTDPSQVSHGTTSGFLKL 471
Db 227 gvfelqpgasvfnvntdpsqvhgtqftsfqllkl 261

RESULT 2

AAR3959
ID AAR3959 standard; Protein; 261 AA.

AC AAR3959;

XX DT 11-JAN-1995 (first entry)

Human CD40-L type II transmembrane protein.

XX KW Leucine zipper; trimerization; trimeric CD40-L; fusion protein; hetero-oligomer; homo-oligomer; type II transmembrane protein; soluble CD40-L; tumour necrosis factor family.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Region 50..261

FT /label= "soluble CD40-L"

WO9410308-A.

XX PD 11-MAY-1994.

XX PF 20-OCT-1993; 93W0-US10034.

XX PR 23-OCT-1992; 93US-0969703.

XX PR 13-AUG-1993; 93US-0107353.

PA (IMMUNEX) IMMUNEX CORP.

XX

PI Spt1995 MK, Srinivasan S;

XX

DR WPI: 1994-167465/20.

XX

PS N-PSDB; AAR3959.

XX

PT Prepn. of soluble oligomeric mammalian proteins - using host cells to express a fusion protein comprising a leucine zipper domain and a heterologous mammalian protein

XX

Example 1; Page 24; 35pp; English.

XX

CC A DNA fragment encoding the extracellular (soluble) region of human CD40-L was ligated to a synthetic oligonucleotide sequence coding CC for a leader peptide, a 33 amino acid leucine zipper sequence (AAR3968) and the Flag (RTM) linker sequence. Cells expressing the CC fusion construct are grown to accumulate oligomeric, soluble CD40-L CC in the supernatant. The leucine zipper sequence spontaneously CC trimers in solution and fusion proteins comprising the CC sequence fused to a heterologous mammalian protein also form CC oligomers.

XX

Query Match 44.6%; Score 1110; DB 15; Length 261;

Best Local Similarity 100.0%; Pred. No. 1.3e-73; Indels 0; Gaps 0;
Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;QY 257 HRRDKIEDRNHLRDFVFMKTIQRCNGERSLSLNEEIKSQFEGFVKDMLNNEETK 316
Db 47 hrrdkiedrnhlrdedfvfmktiqrcngerslslnceekksqfegfvkdlmkeetk 106
QY 317 KENFEMOKGDQDQPIQIAHVVISASSKTSVQIWAEGQYTMNNLVTLENGKQLTVKQ 376
Db 107 kensfemqkdgqpqiahvvisassktsvqiwaeqytmnnlvtlengqqltvkq 166
QY 377 GLYYIAYQVFCNREASSQAPIASLCKSICRERFLILRAANTHSSAKPCGQOIHG 436
Db 167 glyiyiayqvtfcnreassqapiaslickspqrerflilaanttssakpcggqshig 226
QY 437 GVFELQPGASVFNWNTDPSQVSHGTTSGFLKL 471
Db 227 gvfelqpgasvfnvntdpsqvhgtqftsfqllkl 261

RESULT 3

AAR57469
ID AAR57469 standard; Protein; 261 AA.

AC AAR57469;

XX DT 23-MAR-1995 (first entry)

XX DE CD40 ligand.

XX KW Probe; Primer; PCR; amplify; polymerase chain reaction; detection; KW mutation; CD40 ligand gene; IgM; ss.

OS Synthetic.

XX KW Probe; Primer; PCR; amplify; polymerase chain reaction; detection; KW mutation; CD40 ligand gene; IgM; ss.

XX OS Synthetic.

XX PN WO9417196-A.

XX PD 04-AUG-1994.

XX PF 21-JAN-1994; 94W0-US00786.

XX PR 22-JAN-1993; 93US-0009258.

XX PR 20-JAN-1994; 94US-0184422.

XX PA (IMMUNEX) IMMUNEX CORP.

XX

PI	Armitage RJ, Davison BL, Fanslow WC, Renshaw BR;	PD	09-NOV-1995.
PT	Sp19gs MK, Widmer MB;	XX	
XX		XX	
DR	WPI: 1994-264109/32.	PF	28-APR-1995; 95W0-US05448.
DR	N-PSDB: AA067123.	XX	
PT	Method for detecting mutation in CD 40 ligand gene - comprises amplification of nucleic acid, and mutational analysis	PR	28-APR-1994; 94US-0234580.
XX		XX	
PS	Disclosure: Page 22-24; 38pp; English.	DR	(BOEH) BOEHRINGER INGELHEIM PHARM INC.
XX		XX	
CC	This sequence is encoded by the CD40 ligand gene. Mutations within the CD40 ligand gene were identified by the method of the invention. The method comprises isolating DNA from an individual and selectively amplifying the isolated DNA derived from the CD40 ligand gene. The amplification product is then analysed to determine if there is a mutation present and determining if a protein expressed from the ligand gene will bind CD40. The detection of mutations in the CD40 ligand gene allows subsequent treatment of a syndrome resulting in elevated levels of serum IgM and diminished levels of other Ig isotypes, due to mutation in the CD40 ligand gene. ie. X-linked hyper IgM syndrome.	PA	Castle BE, Kehry M;
CC		XX	DR
CC	Sequence 261 AA;	PS	WPI: 1995-393038/50.
CC		XX	N-PSDB: AAT05763.
Query Match	44.6%; Score 110; DB 15; Length 261;	CC	Human high-density, membrane bound CD40 ligand - for stimulating the
Best Local Similarity	100.0%; Pred. No. 1. 3e-73; Mismatches 0; Indels 0; Gaps 0;	CC	type 2 membrane glycoprotein expressed in activated T-cells. It is a member of the tumour necrosis factor receptor family. hmbcD40
Matches	215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	CC	induces long-term proliferation of B-cells in culture. These
Db	257 HRRDLKIEDERNHEDEVFMKTIQRCNIGERSLSLNLCEERKSQFEGFVKDMLNLNEETK 316	CC	proliferating B-cells can be induced to differentiate into antibody-
Db	47 hrrdlkiedernhedfvfmktigcrgerslslceeksqfegfvkdmnlneetk 106	CC	producing cells. Proliferation and differentiation of B-cells
Qy	317 KENSFEMKGDKDORQIAHVTSEASSKTTSVIWAEKYTMNSNLVTLENGKQLTWKQ 376	CC	antigen preferentially proliferates or selects differentiating B-cells
Db	107 kensfemkgdkdgpqiahviaseasskttsviwaekytmnsnlvtlengkqltwkq 166	CC	is obt'd. by incorporating encoding DNA (AAT05763) into a baculovirus
Qy	377 GLYYIYAOVTRENSNREASSQAPPIASLICKLSPGRFERTLRLAANTHSSAKPCGQOSHLG 436	CC	vector that is then used to transfect Sf9 insect cells.
Db	167 qiyiyiyqytfcsnreasqapflasikspgrfertlrlaanthssakpcgqoshlq 226	XX	Sequence 261 AA;
Qy	437 GVFELQPGASVFNVTDPSQVSGTGTGTFSGLUKL 471	Qy	Query Match
Db	227 gvfelqpgasvfnvtdpqsvqshgtgtsgflkl 261	Db	44.6%; Score 110; DB 16; Length 261;
RESULT	4	Qy	Best Local Similarity
DE	AAR85486	Db	100.0%; Pred. No. 1. 3e-73; Mismatches 0; Indels 0; Gaps 0;
ID	AAR85486 standard; Protein: 261 AA.	Qy	Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX		Db	257 HRRDLKIEDERNHEDEVFMKTIQRCNIGERSLSLNLCEERKSQFEGFVKDMLNLNEETK 316
AC	AAR85486;	Db	47 hrrdlkiedernhedfvfmktigcrgerslslceeksqfegfvkdmnlneetk 106
XX		Qy	317 KENSFEMKGDKDORQIAHVTSEASSKTTSVIWAEKYTMNSNLVTLENGKQLTWKQ 376
DT	18-MAR-1996 (first entry)	Db	107 kensfemkgdkdgpqiahviaseasskttsviwaekytmnsnlvtlengkqltwkq 166
XX		Qy	377 GLYYIYAOVTFCNREASSQAPPIASLICKLSPGRFERTLRLAANTHSSAKPCGQOSHLG 436
DE	Human CD40 ligand	Db	167 qiyiyiyqytfcsnreasqapflasikspgrfertlrlaanthssakpcgqoshlq 226
XX		Qy	437 GVFELQPGASVFNVTDPSQVSGTGTGTFSGLUKL 471
KW	High density membrane-bound CD40 ligand; B-lymphocyte; B-cell; differentiation; proliferation; baculovirus; Spodoptera frugiperda; Sf9; insect cell culture; tumour necrosis factor receptor; OS Homo sapiens.	Db	227 gvfelqpgasvfnvtdpqsvqshgtgtsgflkl 261
Key	Location/Qualifiers	RESULT	5
FT	Modified-site 6..8	ID	AAT34669
FT	/label= N-glycosylation_site	ID	AAT34669 standard; Protein: 261 AA.
FT	Domain 23..46	XX	AAT34669;
FT	/label= Transmembrane_domain 239..241	XX	16-FEB-1998 (first entry)
FT	Modified-site /label= N-glycosylation_site	XX	
XX	W09529935-A1.	DE	Amino acid sequence for CD40L, a novel cytokine ligand for CD40.
XX		XX	Cytokine; activated CD4+ T cell; CD40; CD40L; monoclonal antibody; AIDS; neoplastic cell proliferation; B-cell lymphoma; immune deficiency; AIDS; melanoma; carcinoma.
OS	Home sapiens.	XX	
PN	US5674492-A.	XX	
XX		PD	07-OCT-1997.

XX	PP	21-DEC-1994;	94US-0360923.
XX	PR	23-DEC-1993;	93US-0172664.
XX	PR	(USSH) IMMUNEX CORP.	
XX	PI	AMRITAGE RJ, Fanslow WC, Longo DL, Murphy WJ;	
XX	DR	WPI: 1991-502273/46.	
XX	DR	N-PSDB; RAT93782.	
XX	PT	Treating or preventing neoplastic disease associated with CD40-expressing cells - particularly B-cell lymphoma, by administration of CD40-binding protein, preferably antibody or soluble CD40-ligand	
XX	PS	Disclosure; Columns 21-24; 21pp; English.	
CC	CC	The present sequence represents the amino acid sequence of a novel cytokine ligand for CD40 called CD40L. CD40L is a type II membrane polypeptide, which is expressed by activated CD4+ T cells. It has an extracellular region at its C-terminus, and an intracellular region at its N-terminus. The protein can be used to produce monoclonal antibodies, which in turn bind to CD40-expressing cells. This inhibits binding of soluble CD40 to its ligand CD40L. The monoclonal antibody against CD40L is used to inhibit proliferation of neoplastic cells, and is particularly useful in treating B cell lymphoma (e.g. where induced after transplants or in other cases of immune deficiency such as AIDS), and also melanoma or carcinoma. Since the monoclonal antibodies inhibit neoplastic cells directly, they may not need to be coupled to a toxin or radioisotope, avoiding toxic effects on normal B cells.	
CC	CC	Sequence 261 AA;	
CC	CC	Query Match 44.6%; Score 1110; DB 18; Length 261; Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	257	HRRLDKIEDERNLHEDFVFMKTIQRTNGERSLSSLNCEETKSQFEGFKDMLNKEETK 316	
Db	47	hrrlkiedernlhedfvfmktiqrgtngerslsllnceeksqfegfkdmnlneektk 106	
QY	317	KENSFEMQKGDNQPIQRAHVISSEASRSTSQTSLVQWAEGKTYTMSNNLVTENGKLVKQ 376	
Db	107	kensfemqkgdqnpqiaahvisseasrstsqttslvqwaekgytmsnnlvtengkqtkvq 166	
QY	377	GIYYIQAQVFCNSRRAASSQAPFTIASLICKSPGRERILRAANTHSSAKPCGQSIHG 436	
Db	167	giyyiqaqvtfsnreasqapflasickspgrerilraanthssakpcggqsihg 226	
QY	437	GYFELQPGASVFNVNTDPDSQVSHGCTGFTSGLKL 471	
Db	227	gyfelqpgasvfnvntdpdsqvhgtfsqflkl 261	
RESULT	6		
ID	AAW1751	AAW1751 standard; Protein; 261 AA.	
XX	AC	AAW1751;	
XX	DT	08-DEC-1998 (first entry)	
XX	DE	Human CD40 ligand.	
XX	KW	Human; CD40 ligand; TNF receptor family; activated T cell; type 2 membrane glycoprotein; cell proliferation; differentiation; cell.	
XX	OS	Homo sapiens.	
XX	PP	21-DEC-1994;	94US-0360923.
XX	PR	23-DEC-1993;	93US-0172664.
XX	PD	06-OCT-1998.	
XX	PT	28-APR-1995;	95US-0431055.
XX	PR	28-APR-1995;	95US-0431055.
XX	DR	28-APR-1994;	94US-0234580.
XX	PA	(BOEH) BOEHRINGER INGELHEIM PHARM INC.	
XX	PT	Increased proliferation of B cells in culture - by incubating them in the presence of membrane-bound CD40 ligand	
XX	PS	Example 1; Fig 1; 37pp; English.	
CC	CC	The present sequence represents human CD40 ligand which is used in the method of the invention. The method has been developed for proliferating B cells to increase their number at least 100-fold. The method comprises: (a) providing high density membrane bound CD40 ligand; and (b) culturing one or more B cells in the presence of this ligand. The culture results in a proliferation in the number of B cells of at least 100 fold. Also described is a method as above where the B cells are induced to differentiate into antibody-producing cells in the presence of one or more cytokines. The method can be used for stimulating B-cell proliferation in vitro or in vivo, e.g. for treating conditions in which B-cell proliferation and activation is suppressed. Eight rounds of division over six days can be achieved, corresponding to a 256-fold increase in cell numbers, which is a vast increase compared to previous culturing methods.	
CC	CC	Sequence 261 AA;	
QY	257	HRRLDKIEDERNLHEDFVFMKTIQRTNGERSLSSLNCEETKSQFEGFKDMLNKEETK 316	
Db	47	hrrlkiedernlhedfvfmktiqrgtngerslsllnceeksqfegfkdmnlneektk 106	
QY	317	KENSFEMQKGDNQPIQRAHVISSEASRSTSQTSLVQWAEGKTYTMSNNLVTENGKLVKQ 376	
Db	107	kensfemqkgdqnpqiaahvisseasrstsqttslvqwaekgytmsnnlvtengkqtkvq 166	
QY	377	GIYYIQAQVFCNSRRAASSQAPFTIASLICKSPGRERILRAANTHSSAKPCGQSIHG 436	
Db	167	giyyiqaqvtfsnreasqapflasickspgrerilraanthssakpcggqsihg 226	
QY	437	GYFELQPGASVFNVNTDPDSQVSHGCTGFTSGLKL 471	
Db	227	gyfelqpgasvfnvntdpdsqvhgtfsqflkl 261	
RESULT	7		
XX	AAW41178	AAW41178 standard; Protein; 261 AA.	
XX	AC	AAW41178;	
XX	DT	13-MAY-1998 (first entry)	
XX	DE	CD40 ligand.	
XX	KW	Leucine zipper; fusion protein production; soluble oligomeric protein; heterologous mammalian type II transmembrane protein; activated T cell.	
XX	OS	Homo sapiens.	

AC AAY25155;
 XX
 DT 27-AUG-1999 (first entry)
 XX DE Human soluble CD40L protein.
 XX
 CD40L; human; soluble; inhibitor; immune response; T-cell function; alloimmunization; immunotherapy; antibody response; model; PBL; Th cell; peripheral blood lymphocytes; therapy; B cell; APC; CD40; T-helper cell; cytokine secretion; transfusion-induced alloimmune response; anti-HA; systemic lupus erythematosus; SLE; Sjögren's syndrome; Raynaud's syndrome; scleroderma myositis; type 1 diabetes; arthritis; rheumatoid arthritis; inflammatory bowel disease; uveitis; Sjögren's syndrome; Raynaud's syndrome; multiple sclerosis; idiopathic thrombocytopenic purpura; neuroprotective; graft vs host disease; dermatological; immunosuppressive; antidiabetic; antiinflammatory; antirheumatic; antiarthritic; antiallergic.
 XX OS Homo sapiens.
 XX PN WO927948-A2.
 XX
 10-JUN-1999.
 PP 27-NOV-1998; 98WO-CA01105.
 XX PR 28-NOV-1997; 97CA-223325.
 XX PA (CABL-) CANADIAN BLOOD SERVICES.
 XX DE Human CD40 ligand.
 PI CROW JR, Freedman J, Lazarus RH;
 XX DR WPI: 1999-404800/34.
 XX PT Use of recombinant human CD40L protein for treating immune diseases
 XX
 Claim 2: Page 41-42; 42pp; English.
 CC This invention describes a novel use of a soluble recombinant human CD40L (rhCD40L) protein or protein fragment (containing the active binding site with CD40) for inhibiting an immune response and T-cell function. The invention also describes (1) an immunodeficient mouse model of human alloimmunization for testing in vivo effects of an immunotherapy or inhibition of a human antibody response, where the mouse model is an immunodeficient mouse reconstituted with human peripheral blood lymphocytes (PBL) from donors and (2) A method for inhibiting an immune response, or T-cell function, in a patient by administering a therapeutic amount of rhCD40L or functional fragment. The soluble recombinant CD40L active fragment competes B cell-(or antigen presenting cell (APC)-) CD40 interaction with CD40L on the Th (T-helper) cells which disallows the Th cell to be activated to secrete cytokines (such as Th2 cytokines) which thus reduce the transfusion-induced alloimmune response. The rhCD40L or its fragment is useful for inhibiting a human anti-HLA alloimmune response. Where T-cell function is inhibited the following diseases can be treated or prevented: systemic lupus erythematosus (SLE), Sjögren's syndrome, scleroderma myositis, Raynaud's syndrome, type 1 diabetes, arthritis, and rheumatoid arthritis; inflammatory bowel disease, uveitis, myasthenia gravis, multiple sclerosis, idiopathic thrombocytopenic purpura and graft vs host disease, and allergies dependent on T-cells. The development of the model will allow the development of new strategies for the modulation of human alloimmunization to blood cell antigens.
 CC Administration of a recombinant 18kDa CD154 molecule can inhibit an alloimmune response. This molecule may have good therapeutic potential.
 CC The invention has dermatological, immunosuppressive, antiinflammatory, antirheumatic, antiarthritic, antidiabetic, neuroprotective and antiallergic activity.
 CC
 XX Sequence 261 AA;

Query Match 44.6%; Score 110; DB 20; Length 261;

Best Local Similarity 100.0%; Pred. No. 1 3e-73; Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0; QY 257 HRRDLKIEDERNLHEDFVFMKTIORCNGERSLSLNECTKSQFGFVNDIMLNKEETK 316 Db 47 hrrldkiedernlhedfvmktlqrcngerslslneceiksqfegfvkdmLnkeekth 106 QY 317 KENSFEMQGDKDNPQIAHVISSEASSTTSYLOASKGYVMSNNVTLENGQLVKRO 376 Db 107 kensfengkdgdpqiaahnvisseassttsylqwaekgqyvmsnnvltengqtvkq 166 QY 377 GLVYTQAVTFCNSREASSQAFIASCLMSPGRDRILRANHSSAKPCGCGOOSHG 436 Db 167 gVYIYIeVqTfcsnreassqafiasclmspgrferrllraanthssakpcggqshg 226 QY 437 GYFELORGASVENVNTPSVPQSHGTCFTSCLKL 471 Db 227 gVfelpqgasvlnvntpsvshgtgftsfyjkl 261

RESULT 10
 ID AAY96993
 XX ID AAY96993 standard; Protein; 261 AA.
 XX AC AAY96993;
 XX DT 31-OCT-2000 (first entry)
 XX DE Human CD40 ligand.
 XX
 XX KW CD40 ligand; CD40; T cell; T cell receptor; rearrangement; maturation; cell death inhibition; stress-induced; immunosuppressive; anti-thyroid; anti-inflammation; anti-diabetic; anti-rheumatic; anti-anaemic; ophthalmological; anti-poriactic; nephrotrophic; hepatotropic; virucide; dermatological; cytostatic.
 XX OS Homo sapiens.
 XX PN WO200039283-A1.
 XX PD 06-JUL-2000.
 XX PP 22-DEC-1999; 99WO-US30930.
 XX PR 29-DEC-1998; 98US-014106.
 XX PA (UVM-) UNIV VERMONT & STATE AGRIC COLLEGE.
 XX PI Newell MK, Wagner D, Newell E;
 XX DR WPI: 2000-452387/39.
 XX DR N-PSDB: ANA51745.

PT Inducing T cell receptor gene rearrangement for treating autoimmune diseases comprises contacting T cells with a CD40-binding agent Disclosure; Page 47; 50pp; English.

CC CD40 engagement on T cells can be used to induce T cell receptor rearrangement and enhance T cell affinity for a particular antigen. CD40 engagement can be bought about by contacting CD40 with a CD40-binding agent, e.g. human CD40 ligand. The CD40-binding agents can also be used in methods for promoting T cell maturation, inhibiting T cell receptor rearrangement, inhibiting environmental stress-induced cell death, altering the specificity of a T cell towards an antigen, inducing T cell reactivity towards an antigen or enhancing environmental stress-induced cell death (all claimed). T cell affinity maturation towards a specific antigen can be inhibited, especially for a self-antigen in an autoimmune disease, which includes rheumatoid arthritis, uveitis, insulin-dependent diabetes mellitus, haemolytic anaemias, rheumatic fever, Crohn's disease, Guillain-Barre syndrome, psoriasis, thyroiditis, Grave's disease, myasthenia gravis, glomerulonephritis, autoimmune hepatitis or systemic lupus erythematosus. Inducing environmental stress-induced T cell death

CC is carried out in a cancerous T cell or a self-reactive T cell where the environmental stress is a chemotherapeutic agent (claimed).
 CC environmental stress is a chemotherapeutic agent (claimed).
 XX Sequence 261 AA;
 SQ

Query Match 44.6%; Score 1110; DB 21; Length 261;
 Best Local Similarity 100.0%; Pred. No. 1 3e-23; Mismatches 0; Indels 0; Gaps 0;
 Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 AC

Oy 257 HRLDKTIEDERNHEDVFMTKTCRNCNGERSLSLNLCEETKSQFEGFKDMLNKEETK 316
 Db 47 hrrldkldernhdedfvfmktlqrcntgterslslinceeklsqfedfvkdlmlnkeek 106

Qy 317 KENSFEMQKGDNQPIQIAHWYSEASSKITSVQWAEGKTYMSNNLYTLENGKQTVKQ 376
 Db 107 kensfemqkgdqnpqiahwseassktsvqlwaekgytunnnvlengkqtvkq 166

Db 377 GLYYIYAQVTCGSNREASSQAPTIASCLKSPCRFERILLRAANTHSSAKPCGQSQTHLG 436
 167 glyiyiyaqtfcsnreassqapfiaslkspcrferillraanthssakpcgqqsihlg 226

Qy 437 GVELQPCASVFNVTPOSGHGTTSFGKLK 471
 Db 227 gvelqpgasvfnvntdpqsvngtgftsglk 261

RESULT 11
 AAB67612
 ID AAB67612 standard; Protein: 261 AA.
 AC
 XX
 DT 29-MAY-2001 (first entry)
 DE Amino acid sequence of human gp39 protein, a CD40 ligand.
 XX
 KW gp39; CD40 ligand; osteoblast cell death; apoptosis; bone loss;
 KW osteoporosis; inflammatory arthritis; esterom loss;
 KW ovariectomy; historectomy; lupus nephritis; Takayasu's arteritis;
 KW Wegener's granulomatosis; nephritis; myositis; scleroderma;
 KW thrombocytopenia; asthma; lung disease; cancer.
 XX Homo sapiens.
 OS
 XX
 PN WO200116180-A2.
 PD 08-MAR-2001.
 XX
 XX
 PF 24-AUG-2000; 2000WO-US23276.
 XX
 PR 27-AUG-1999; 99US-0151250.
 XX
 PA (TEXA) UNIV TEXAS SYSTEM.
 XX
 PI Ahuja SS, Bonewald LF;
 XX
 DR WPI; 2001-169007/17.
 DR N-PSDB; AAF55539.

XX
 PT CD40 agonist containing composition, used to reduce bone cell death or apoptosis associated with osteoporosis, osteonclerosis and inflammatory arthritis -
 XX
 PS Claim 3: Page 113-114; 118pp; English.

XX
 CC The present sequence represents a gp39 protein. It is a CD40 ligand.
 CC CD40 ligands are used for reducing osteoblast cell death or apoptosis,
 CC and for treating or preventing bone loss in animals, preferably humans,
 CC at risk of, or undergoing, bone loss. The bone loss is associated with
 CC osteoporosis, osteonecrosis, inflammatory arthritis, post-menopausal
 CC osteoporosis, estrogen loss due to ovariectomy, total historectomy,
 CC lupus nephritis, Takayasu's arteritis, Wegener's granulomatosis,

CC anti-glomerular basement membrane nephritis, myositis, scleroderma,
 CC idiopathic autoimmune thrombocytopenia, asthma, a chronic obstructive
 CC lung disease, nephrotic/nephritic syndrome, or cancer. They may also be
 CC used to treat or prevent bone loss in a subject undergoing, or scheduled
 CC for, an organ or bone marrow transplant.
 XX Sequence 261 AA;

Query Match 44.6%; Score 1110; DB 22; Length 261;
 Best Local Similarity 100.0%; Pred. No. 1 3e-73; Mismatches 0; Indels 0; Gaps 0;
 Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 AC

Oy 257 HRLDKTIEDERNHEDVFMTKTCRNCNGERSLSLNLCEETKSQFEGFKDMLNKEETK 316
 Db 47 hrrldkldernhdedfvfmktlqrcntgterslslinceeklsqfedfvkdlmlnkeek 106

Qy 317 KENSFEMQKGDNQPIQIAHWYSEASSKITSVQWAEGKTYMSNNLYTLENGKQTVKQ 376
 Db 107 kensfemqkgdqnpqiahwseassktsvqlwaekgytunnnvlengkqtvkq 166

Db 377 GLYYIYAQVTCGSNREASSQAPTIASCLKSPCRFERILLRAANTHSSAKPCGQSQTHLG 436
 167 glyiyiyaqtfcsnreassqapfiaslkspcrferillraanthssakpcgqqsihlg 226

Qy 437 GVELQPCASVFNVTPOSGHGTTSFGKLK 471
 Db 167 glyiyiyaqtfcsnreassqapfiaslkspcrferillraanthssakpcgqqsihlg 226

Db 227 gvelqpgasvfnvntdpqsvngtgftsglk 261

RESULT 12
 AAB37806
 ID AAB37806 standard; Protein: 261 AA.
 AC
 XX
 DT 23-FEB-2001 (first entry)
 DE Human CD40 ligand sequence.
 XX
 KW Human; CD40 ligand; antiviral; anti-HIV; CD40 stimulant;
 KW human immunodeficiency virus; HIV; cancer; organ transplantation.
 XX
 OS Homo sapiens.
 XX
 PN WO200066155-A1.
 PD 09-NOV-2000.
 XX
 XX
 PF 28-APR-2000; 2000WO-US11734.
 XX
 PR 30-APR-1999; 99US-0131730.
 XX
 PA (LUDL-) LA JOLLA INST ALLERGY & IMMUNOLOGY.
 XX
 PI Sarawar SR, Schoenberger SP;
 XX
 DR WPI; 2001-024740/03.
 PT Method for preventing latent virus reactivation or controlling virus
 PT replication by administration of antibodies or similar reagents which
 PT stimulate the CD40 molecule -
 XX
 PS Disclosure; Page 16; 36pp; English.

XX
 CC The present sequence is a human CD40 ligand. It may be used in a method
 CC for preventing latent virus reactivation or controlling virus
 CC replication. The method comprises administering a composition capable
 CC of binding to a cell membrane expressed CD40, where the binding of the
 CC composition to the CD40 on the surface of the cell generates a
 CC stimulatory signal to the cell. The method is useful for treating
 CC immunocompromised patients, e.g. infected with human immunodeficiency
 CC virus (HIV), or immunosuppressed patients, e.g. due to infections or

Tue Jul 3 16:26:03 2001

us-09-454-223-2.rag

GenCore version 4.5
 Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model
 Run on: July 3, 2001, 16:13:43 : Search time 17.98 Seconds
 (without alignments)
 527.704 Million cell updates/sec

Title: US-09-454-223-2
 Perfect score: 2487
 Sequence: 1 MFLPLSLMVLVLVQPLGNLGR..... TDSQSVSHTGTFMSFLGLKL 471
 Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5 .
 searched: 193259 seqs, 20144639 residues

Total number of hits satisfying chosen parameters: 193259
 Minimum Match 0%
 Maximum Match 100%
 Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : Issued_Parents_AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A__COMB.pep:*
- 2: /cgn2_6/ptodata/2/1aa/5B__COMB.pep:*
- 3: /cgn2_6/ptodata/2/1aa/6A__COMB.pep:*
- 4: /cgn2_6/ptodata/2/1aa/6B__COMB.pep:*
- 5: /cgn2_6/ptodata/2/1aa/PCTUS__COMB.pep:*
- 6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1110	44.6	261	1 US-07-940-605A-2	Sequence 2, Appl1
2	1110	44.6	261	1 US-08-189-422-8	Sequence 8, Appl1
3	1110	44.6	261	1 US-08-360-923A-2	Sequence 2, Appl1
4	1110	44.6	261	1 US-08-446-922-4	Sequence 4, Appl1
5	1110	44.6	261	2 US-08-431-055-4	Sequence 4, Appl1
6	1110	44.6	261	2 US-08-690-096-2	Sequence 2, Appl1
7	1110	44.6	261	2 US-09-249-189-12	Sequence 12, Appl1
8	1110	44.6	261	2 US-08-489-624N-12	Sequence 12, Appl1
9	1110	44.6	261	2 US-08-477-733B-12	Sequence 12, Appl1
10	1110	44.6	261	3 US-08-767-995-2	Sequence 2, Appl1
11	1110	44.6	261	3 US-09-913A-12	Sequence 12, Appl1
12	1110	44.6	261	3 US-09-589-771B-8	Sequence 8, Appl1
13	1110	44.6	261	5 PCT-US93-10034-4	Sequence 4, Appl1
14	1094.5	44.0	273	1 US-09-465-922-11	Sequence 11, Appl1
15	1094.5	44.0	273	2 US-09-249-189-21	Sequence 12, Appl1
16	1094.5	44.0	273	2 US-09-484-624A-21	Sequence 2, Appl1
17	1094.5	44.0	273	2 US-09-477-733B-21	Sequence 2, Appl1
18	1094.5	44.0	273	3 US-09-088-913A-21	Sequence 2, Appl1
19	1089	43.8	473	2 US-09-249-189-16	Sequence 21, Appl1
20	1089	43.8	473	2 US-09-484-624A-16	Sequence 21, Appl1
21	1089	43.8	473	2 US-08-477-733B-16	Sequence 16, Appl1
22	1089	43.8	473	3 US-09-088-913A-16	Sequence 16, Appl1
23	831	33.4	260	1 US-08-446-922-6	Sequence 6, Appl1
24	831	33.4	260	2 US-09-249-189-2	Sequence 2, Appl1
25	831	33.4	260	2 US-09-484-624A-2	Sequence 2, Appl1
26	831	33.4	260	2 US-08-477-733B-2	Sequence 2, Appl1
27	831	33.4	260	3 US-09-088-913A-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1
 US-07-940-605A-2
 Sequence 2, Application US/07940605A
 Patent No. 5540925

GENERAL INFORMATION:

APPLICANT: ARUFOO, ALEJANDRO
 APPLICANT: HOLLENBAUGH, DIANE
 APPLICANT: LEDBETTER, JEFFREY A.

TITLE OF INVENTION: SOLUBLE LIGANDS FOR CD40
 NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADRESSEE: Pennie & Edmonds
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/940,605A
 FILING DATE: 04-SEP-1992
 CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: MISTOCK, S. Leslie
 REGISTRATION NUMBER: 18,872
 REFERENCE/DOCKET NUMBER: 5624-184

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212 790-9090
 TELEFAX: 212 869-8864/9741
 TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 261 amino acids
 TYPE: amino acid

TOPOLOGY: linear
 MOLECULE TYPE: protein

US-07-940-605A-2

Query Match 44.6%; Score 1110; DB 1; Length 261;
 Best Local Similarity 100%; pred. No. 3.5e-79;
 Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 257 HRRRLKIEDERNLHDSPFVFMKTIQRCNTGERSLSLJNCEEKSOFEGFVFDIMLNKEETK 316
 DB 47 HRRRLKIEDERNLHDSPFVFMKTIQRCNTGERSLSLJNCEEKSOFEGFVFDIMLNKEETK 106

QY 317 KENSFEMOKDQDNPQIAAHVISEASSKTTSVLQWAERGKYYTMSNNLVLENGKOLTVKRO 376
 Db 107 KENSFEMOKDQDNPQIAAHVISEASSKTTSVLQWAERGKYYTMSNNLVLENGKOLTVKRO 166
 QY 377 GLYIYIAQVTCNSREASSOAPFTASICKSLSGRFERILLRAANTHSSAKPCGQOSIHLG 436
 Db 167 GLYIYIAQVTCNSREASSOAPFTASICKSLSGRFERILLRAANTHSSAKPCGQOSIHLG 226
 QY 437 GVFLQPGASVFNVTDPSONSHGTGFTSGLLKL 471
 Db 227 GVFLQPGASVFNVTDPSONSHGTGFTSGLLKL 261

RESULT 2
 US-08-184-422-8
 ; Sequence 8, Application US/08184422
 ; Patent No. 5,65321
 ; GENERAL INFORMATION:
 ; APPLICANT: ARMITAGE, RICHARD
 ; APPLICANT: DAVISON, BARRY
 ; APPLICANT: FANSLOW, WILLIAM
 ; APPLICANT: RENSHAW, BLAIR
 ; APPLICANT: SPRIGGS, MELANIE
 ; APPLICANT: WIDMER, MICHAEL
 ; TITLE OF INVENTION: DETECTION AND TREATMENT OF MUTATIONS
 ; TITLE OF INVENTION: IN A C040 LIGAND GENE
 ; NUMBER OF SEQUENCES: 16
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: IMMUNEX CORPORATION
 ; STREET: 51 UNIVERSITY STREET
 ; CITY: SEATTLE
 ; STATE: WASHINGTON
 ; COUNTRY: USA
 ; ZIP: 98101
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: Apple Macintosh
 ; OPERATING SYSTEM: Apple Operating System 7.1
 ; SOFTWARE: MS Word for Apple 5.1, Version a
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/184,422
 ; FILING DATE: 01/22/93
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: PERKINS, PATRICIA ANNE
 ; REGISTRATION NUMBER: 34,693
 ; REFERENCE/DOCKET NUMBER: 2810-A
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (206)587-0430
 ; TELEFAX: 2065870506
 ; INFORMATION FOR SEQ ID NO: 8:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 261 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-184-422-8

RESULT 3
 US-08-360-923A-2
 ; Sequence 2, Application US/08360923A
 ; Patent No. 5,674492
 ; GENERAL INFORMATION:
 ; APPLICANT: ARMITAGE, RICHARD
 ; APPLICANT: FANSLOW, WILLIAM
 ; APPLICANT: LONGO, DAN L.
 ; APPLICANT: MURPHY, WILLIAM
 ; TITLE OF INVENTION: METHOD OF PREVENTING OR TREATING
 ; TITLE OF INVENTION: DISASE CHARACTERIZED BY NEOPLASITC CELLS
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: IMMUNEX CORPORATION
 ; STREET: 51 UNIVERSITY STREET
 ; CITY: SEATTLE
 ; STATE: WASHINGTON
 ; COUNTRY: USA
 ; ZIP: 98101
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: Macintosh
 ; OPERATING SYSTEM: Apple Macintosh System 7.1
 ; SOFTWARE: Microsoft Word for Macintosh, Version 4.5.1a
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/360,923A
 ; FILING DATE: December 21, 1994
 ; CLASSIFICATION: 424
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: USN 08/172,664
 ; FILING DATE: December 23, 1993
 ; CLASSIFICATION: 424
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Perkins, Patricia A.
 ; REGISTRATION NUMBER: 34,693
 ; REFERENCE/DOCKET NUMBER: 2810-A
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (206)233-0644
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 261 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-360-923A-2

Query Match 44.6%; Score 1110; DB 1; Length 261;
 Best Local Similarity 100.0%; Pred. No. 3.5e-79; Length 261;
 Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 257 HRLDKIEDERNLHEDFVNPKTQRCNTGERSLSLLNGEETIKSOFEGVKDMLNKEETK 316
 Db 47 HRLDKIEDERNLHEDFVNPKTQRCNTGERSLSLLNGEETIKSOFEGVKDMLNKEETK 106
 QY 317 KENSFEMOKDQDNPQIAAHVISEASSKTTSVLQWAERGKYYTMSNNLVLENGKOLTVKRO 376
 Db 107 KENSFEMOKDQDNPQIAAHVISEASSKTTSVLQWAERGKYYTMSNNLVLENGKOLTVKRO 166
 QY 377 GLYIYIAQVTCNSREASSOAPFTASICKSLSGRFERILLRAANTHSSAKPCGQOSIHLG 436
 Db 167 GLYIYIAQVTCNSREASSOAPFTASICKSLSGRFERILLRAANTHSSAKPCGQOSIHLG 226
 QY 437 GVFLQPGASVFNVTDPSONSHGTGFTSGLLKL 471
 Db 227 GVFLQPGASVFNVTDPSONSHGTGFTSGLLKL 261

Query Match 44.6%; Score 1110; DB 1; Length 261;
 Best Local Similarity 100.0%; Pred. No. 3.5e-79; Length 261;
 Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 257 HRLDKIEDERNLHEDFVNPKTQRCNTGERSLSLLNGEETIKSOFEGVKDMLNKEETK 316
 Db 47 HRLDKIEDERNLHEDFVNPKTQRCNTGERSLSLLNGEETIKSOFEGVKDMLNKEETK 106
 QY 317 KENSFEMOKDQDNPQIAAHVISEASSKTTSVLQWAERGKYYTMSNNLVLENGKOLTVKRO 376
 Db 107 KENSFEMOKDQDNPQIAAHVISEASSKTTSVLQWAERGKYYTMSNNLVLENGKOLTVKRO 166
 QY 377 GLYIYIAQVTCNSREASSOAPFTASICKSLSGRFERILLRAANTHSSAKPCGQOSIHLG 436

Db 167 GLYYIYAOVTFCNSREASSQAPIIASCLKSPLRERFILLRANTHSSAKPCGQOSHLG 226
 Qy 437 GVFELQPGASVNVNTDPSQVSHGTGFTSGFLKL 471
 Db 227 GVFELQPGASVNVNTDPSQVSHGTGFTSGFLKL 261

RESULT 4
 US-08-46-922-4
 ; Sequence 4, Application US/08446922
 ; Patent No. 5716805
 GENERAL INFORMATION:
 APPLICANT: Srinivasan, Subhashini
 TITLE OF INVENTION: Methods of Preparing Soluble, Oligomeric
 NUMBER OF SEQUENCES: 11
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Immunex Corporation
 STREET: 51 University Street
 CITY: Seattle
 STATE: WA
 COUNTRY: USA
 ZIP: 98101

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: Apple Macintosh
 OPERATING SYSTEM: Apple Operating System 7.1
 SOFTWARE: Microsoft Word for Apple, Version 5.1a

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/446,922
 FILING DATE: 08-13-93
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Perkins, Patricia A
 REGISTRATION NUMBER: 34,693
 REFERENCE/DOCKET NUMBER: 1003-A

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: USN 08/107,353
 FILING DATE: 08-13-93
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: MILLMAN, ROBERT A
 REGISTRATION NUMBER: 36,217
 REFERENCE/DOCKET NUMBER: 1011.1030000/RAM

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206) 587-0430
 TELEFAX: (206) 233-0644

INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 261 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-446-922-4

Query Match 44.6%; Score 1110; DB 1; Length 261;
 Best Local Similarity 100.0%; Pred. No. 3; 5e-79;
 Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 257 HRRLDKIEDERNLHEDVFVFKTIIORCNGERSLSLLNCEEIKSOFEGFVKDMLNKEETK 316
 Db 47 HRRLDKIEDERNLHEDVFVFKTIIORCNGERSLSLLNCEEIKSOFEGFVKDMLNKEETK 106
 Qy 317 KENSFEMQKGDNPOIAAHVISEBASSKTTSVLQWAEGKYTMNSNLVLENGKQLTVKRO 376
 Db 167 GLYYIYAOVTFCNSREASSQAPIIASCLKSPLRERFILLRANTHSSAKPCGQOSHLG 226
 Qy 317 KENSFEMQKGDNPOIAAHVISEBASSKTTSVLQWAEGKYTMNSNLVLENGKQLTVKRO 166
 Db 107 KENSFEMQKGDNPOIAAHVISEBASSKTTSVLQWAEGKYTMNSNLVLENGKQLTVKRO 166

Query Match 44.6%; Score 1110; DB 1; Length 261;
 Best Local Similarity 100.0%; Pred. No. 3; 5e-79;
 Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 257 HRRLDKIEDERNLHEDVFVFKTIIORCNGERSLSLLNCEEIKSOFEGFVKDMLNKEETK 316
 Db 47 HRRLDKIEDERNLHEDVFVFKTIIORCNGERSLSLLNCEEIKSOFEGFVKDMLNKEETK 106
 Qy 317 KENSFEMQKGDNPOIAAHVISEBASSKTTSVLQWAEGKYTMNSNLVLENGKQLTVKRO 376
 Db 167 GLYYIYAOVTFCNSREASSQAPIIASCLKSPLRERFILLRANTHSSAKPCGQOSHLG 226
 Qy 437 GVFELQPGASVNVNTDPSQVSHGTGFTSGFLKL 471
 Db 227 GVFELQPGASVNVNTDPSQVSHGTGFTSGFLKL 261

RESULT 5
 US-08-31-055-4
 ; Sequence 4, Application US/08431055
 ; Patent No. 5817516
 GENERAL INFORMATION:
 APPLICANT: KERRY, MERILYN R
 APPLICANT: CASTLE, BRIAN E
 TITLE OF INVENTION: METHODS FOR PROLIFERATING AND
 DIFFERENTIATING B CELLS, AND USES THEREOF
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: STEIN, KESSLER, GOLDSTEIN & FOX
 STREET: 100 NEW YORK AVE. N.W. SUITE 600
 CITY: WASHINGTON
 STATE: D.C.
 ZIP: 20005

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US 08/234,580
 FILING DATE: 28-APR-1994
 CLASSIFICATION: 435
 PRIORITY APPLICATION DATA:
 FILING DATE: 28-APR-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: MILLMAN, ROBERT A
 REGISTRATION NUMBER: 36,217
 REFERENCE/DOCKET NUMBER: 1011.1030000/RAM

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 371-2600
 TELEFAX: (202) 371-2540

INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 261 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-431-055-4

Query Match 44.6%; Score 1110; DB 2; Length 261;
 Best Local Similarity 100.0%; Pred. No. 3; 5e-79;
 Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 257 HRRLDKIEDERNLHEDVFVFKTIIORCNGERSLSLLNCEEIKSOFEGFVKDMLNKEETK 316
 Db 47 HRRLDKIEDERNLHEDVFVFKTIIORCNGERSLSLLNCEEIKSOFEGFVKDMLNKEETK 106
 Qy 317 KENSFEMQKGDNPOIAAHVISEBASSKTTSVLQWAEGKYTMNSNLVLENGKQLTVKRO 376
 Db 167 GLYYIYAOVTFCNSREASSQAPIIASCLKSPLRERFILLRANTHSSAKPCGQOSHLG 226
 Qy 437 GVFELQPGASVNVNTDPSQVSHGTGFTSGFLKL 471
 Db 227 GVFELQPGASVNVNTDPSQVSHGTGFTSGFLKL 261

RESULT 6
 US-08-60-096-2
 ; Sequence 2, Application US/08690096
 ; Patent No. 5945513
 ; GENERAL INFORMATION:
 APPLICANT: ARUFFO, ALEJANDRO

APPLICANT: HOLLOWBAUGH, DIANE
 APPLICANT: LIEBETTER, JEFFREY A.
 TITLE OF INVENTION: SOLUBLE LIGANDS FOR CD40
 NUMBER OF SEQUENCES: 15
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Penne & Edmonds
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10036-2711

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, version #1.25

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/690.096
 FILING DATE: 31-JUL-1996
 CLASSIFICATION: 536
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 07/940,605
 FILING DATE: 04-SEP-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: MISHOCK, S. Leslie
 REFERENCE/DOCKET NUMBER: 18 872
 INFORMATION FOR SEQ ID NO: 2:
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212 790-9090
 TELEFAX: 212 869-8864/8741
 TELEX: 66141 PENNIE
 LENGTH: 261 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-690-096-2

Query Match 44.6%; Score 1110; DB 2; Length 261;
 Best Local Similarity 100.0%; Pred. No. 3.5e-79; Mismatches 0; Indels 0; Gaps 0;
 Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 257 HRRDKIEDERNLHEDFVFMKTIORCNGERSLSLNLCECEBKSQEGFKDILNKEETK 316
 Db 47 HRRDKIEDERNLHEDFVFMKTIORCNGERSLSLNLCECEBKSQEGFKDILNKEETK 106

QY 317 KENSFEMOKGDNPOIAHVTSEASSKTSVLOQWAEGKGYMSNNLVTLENGKQLTVKRO 376
 107 KENSFEMOKGDNPOIAHVTSEASSKTSVLOQWAEGKGYMSNNLVTLENGKQLTVKRO 376
 QY 377 GLYYIYAQVTCNSREASSQAFASLCKSPGFRERILRAANTHSSAKPCGQOOSIHLG 436
 Db 167 GLYYIYAQVTCNSREASSQAFASLCKSPGFRERILRAANTHSSAKPCGQOOSIHLG 226

QY 437 GVFELQPGASVFWVNTDPSQVSHGTTGFTSRLKL 471
 Db 227 GVFELQPGASVFWVNTDPSQVSHGTTGFTSRLKL 261

RESULT 7
 US-08-249-189-12
 ; Sequence 12, Application US/08249189
 ; Patent No. 596194
 ; GENERAL INFORMATION:
 ; APPLICANT: ARMITAGE, RICHARD
 ; APPLICANT: FANSLOW, WILLIAM
 ; APPLICANT: SPRIGGS, MELANIE
 ; APPLICANT: SRIINTIVASAN, SUBHASHINI
 ; APPLICANT: GIBSON, MARYLOU
 ; TITLE OF INVENTION: NOVEL CYTOKINE
 ; NUMBER OF SEQUENCES: 23

Query Match 44.6%; Score 1110; DB 2; Length 261;
 Best Local Similarity 100.0%; Pred. No. 3.5e-79; Mismatches 0; Indels 0; Gaps 0;
 Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 257 HRRDKIEDERNLHEDFVFMKTIORCNGERSLSLNLCECEBKSQEGFKDILNKEETK 316
 Db 47 HRRDKIEDERNLHEDFVFMKTIORCNGERSLSLNLCECEBKSQEGFKDILNKEETK 106

QY 317 KENSFEMOKGDNPOIAHVTSEASSKTSVLOQWAEGKGYMSNNLVTLENGKQLTVKRO 376
 107 KENSFEMOKGDNPOIAHVTSEASSKTSVLOQWAEGKGYMSNNLVTLENGKQLTVKRO 376
 QY 377 GLYYIYAQVTCNSREASSQAFASLCKSPGFRERILRAANTHSSAKPCGQOOSIHLG 436
 Db 167 GLYYIYAQVTCNSREASSQAFASLCKSPGFRERILRAANTHSSAKPCGQOOSIHLG 226

QY 437 GVFELQPGASVFWVNTDPSQVSHGTTGFTSRLKL 471
 QY 377 GLYYIYAQVTCNSREASSQAFASLCKSPGFRERILRAANTHSSAKPCGQOOSIHLG 436
 Db 167 GLYYIYAQVTCNSREASSQAFASLCKSPGFRERILRAANTHSSAKPCGQOOSIHLG 226

QY 437 GVFELQPGASVFWVNTDPSQVSHGTTGFTSRLKL 471
 Db 227 GVFELQPGASVFWVNTDPSQVSHGTTGFTSRLKL 261

RESULT 8
 US-08-484-524-12
 ; Sequence 12, Application US/08484624A
 ; Patent No. 5962406
 ; GENERAL INFORMATION:
 ; APPLICANT: ARMITAGE, RICHARD
 ; APPLICANT: FANSLOW, WILLIAM
 ; APPLICANT: SPRIGGS, MELANIE

CORRESPONDENCE ADDRESS:
 ADDRESSEE: IMMUNEX CORPORATION
 STREET: 51 UNIVERSITY STREET
 CITY: SEATTLE
 STATE: WASHINGTON
 COUNTRY: USA
 ZIP: 98101

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: Apple Macintosh
 OPERATING SYSTEM: Apple Operating System 7.1
 SOFTWARE: Microsoft Word for Apple, version 5.1a
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/249,189
 FILING DATE:
 CLASSIFICATION: 514

PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: 07/969,703
 FILING DATE: October 23, 1992
 CLASSIFICATION: 514

PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: 07/783,707
 FILING DATE: October 25, 1991
 CLASSIFICATION: 514

PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: 07/805,723
 FILING DATE: December 5, 1991
 CLASSIFICATION: 514

PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: 07/783,707
 FILING DATE: October 23, 1992
 CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:
 NAME: Perkins, Patricia A.
 REGISTRATION NUMBER: 34,633
 REFERENCE/DOCKET NUMBER: 2802-C

TELECOMMUNICATION INFORMATION:
 TELEPHONE: 205870430
 TELEFAX: 205870606

INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 261 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-249-189-12

CORRESPONDENCE ADDRESS:
 ADDRESSEE: IMMUNEX CORPORATION
 STREET: 51 UNIVERSITY STREET
 CITY: SEATTLE
 STATE: WASHINGTON
 COUNTRY: USA
 ZIP: 98101

CORRESPONDENCE ADDRESS:
 ADDRESSEE: IMMUNEX CORPORATION
 STREET: 51 UNIVERSITY STREET
 CITY: SEATTLE
 STATE: WASHINGTON
 COUNTRY: USA
 ZIP: 98101

APPLICANT: SRINIVASAN, SUBHASHINI
 APPLICANT: GIBSON, MARYLOU
 APPLICANT: MORRIS, ARVIA E.
 APPLICANT: MCGREW, JEFFERY
 TITLE OF INVENTION: NOVEL CYTOKINE THAT BINDS CD40
 NUMBER OF SEQUENCES: 26
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: IMMUNEX CORPORATION
 STREET: 51 UNIVERSITY STREET
 CITY: SEATTLE
 STATE: WASHINGTON
 COUNTRY: USA
 ZIP: 98101
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: Apple Macintosh
 OPERATING SYSTEM: Apple Operating System 7.5.5
 SOFTWARE: MS Word for Apple Power Macintosh, version 6.0.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/484,624A
 FILING DATE:
 CLASSIFICATION:
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: 08/477,733
 FILING DATE: October 25, 1991
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Perkins, Patricia A.
 REGISTRATION NUMBER: 34,693
 REFERENCE/DOCKET NUMBER: 2802-D
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 2065870430
 TELEFAX: 2065870605
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 261 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-484-624A-12

Query Match 44.6%; Score 1110; DB 2; Length 261;
 Best Local Similarity 100.0%; Pred. No. 3 5e-79; Mismatches 0; Gaps 0; Indels 0;

QY 257 HRRLDKIEDERNHEDVFVFKTIIQRCWIGERSLSLLNCEEIKSQFEGPVKDMLNKEETK 316
 DB 47 HRRLDKIEDERNHEDVFVFKTIIQRCWIGERSLSLLNCEEIKSQFEGPVKDMLNKEETK 106
 ' 317 KENSFEMQKGDNQPOIAHVISSEASSKTTSVLQWAEGYYTMSNNLVLENGKQTVKQ 376
 ' DB 47 HRRLDKIEDERNHEDVFVFKTIIQRCWIGERSLSLLNCEEIKSQFEGPVKDMLNKEETK 106
 ' 377 GLYYIYQVTFCSREASSQAPPIASLICKSPGRFENILRANTHSSAKPCGQOSTHIG 436
 ' DB 167 GLYYIYQVTFCSREASSQAPPIASLICKSPGRFENILRANTHSSAKPCGQOSTHIG 226
 QY 437 GYFELQPGASVVFNTDPSQVSIGTGTSGFLKL 471
 DB 227 GYFELQPGASVVFNTDPSQVSIGTGTSGFLKL 261

Query Match 44.6%; Score 1110; DB 2; Length 261;
 Best Local Similarity 100.0%; Pred. No. 3 5e-79; Mismatches 0; Gaps 0; Indels 0; Gaps 0;

QY 257 HRRLDKIEDERNHEDVFVFKTIIQRCWIGERSLSLLNCEEIKSQFEGPVKDMLNKEETK 316
 ' 317 KENSFEMQKGDNQPOIAHVISSEASSKTTSVLQWAEGYYTMSNNLVLENGKQTVKQ 376
 ' DB 47 HRRLDKIEDERNHEDVFVFKTIIQRCWIGERSLSLLNCEEIKSQFEGPVKDMLNKEETK 106

RESULT 10
 US-08-763-995-2
 Sequence 2, Application US/08763995
 Patent No. 6017527
 GENERAL INFORMATION:
 APPLICANT: MARASKOVSKY, EUGENE
 TITLE OF INVENTION: METHOD OF ACTIVATING DENDRITIC CELLS
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: IMMUNEX CORPORATION
 STREET: 51 UNIVERSITY STREET
 CITY: SEATTLE
 STATE: WASHINGTON
 COUNTRY: USA
 ZIP: 98101
 COMPUTER READABLE FORM:
 MEDIUM TYPE: FLOPPY disk
 COMPUTER: Power Macintosh 7200/90
 OPERATING SYSTEM: Apple Operating System 7.5.5
 SOFTWARE: Microsoft Word for Macintosh, Version #6.0.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/763, 995
 FILING DATE:
 CLASSIFICATION:
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: USN 08/677, 762
 FILING DATE: 10 JUL 1996
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Perkins, Patricia A.
 REGISTRATION NUMBER: 34, 693
 REFERENCE/DOCKET NUMBER: 2845-A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206)587-0430
 TELEFAX: (206)233-0644
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 261 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-763-995-2

Query Match 44 6%; Score 110; DB 3; Length 261;
 Best Local Similarity 100.0%; Pred. No. 3. 5e-79;
 Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 257 HRRLDKIEDERNLHEDFVFMKTIQRCNGERSLSLLNCEEIKSOSFGCFVKDMLNKEETK 316
 QY 317 KENSFEMQKGDNQPOIAAHVIBASSKTSVQWAERGYYMMNNVTLENGKQLTVKRQ 376
 QY 107 KENSFEMQKGDNQPOIAAHVIBASSKTSVQWAERGYYMMNNVTLENGKQLTVKRQ 166
 QY 377 GLYYIYAQVFCNRREASSQAPFTASCLSKSPGRERILLRAANTHSSAKPCGQOOSHG 436
 Db 167 GLYYIYAQVFCNSREASSQAPFTASCLSKSPGRERILLRAANTHSSAKPCGQOOSHG 226
 QY 437 GVFELQPGASVNVNTDPSQVSHGTGFTSFGLKL 261

RESULT 11
 US-09-088-913A-12
 Sequence 12, Application US/09088913A
 Patent No. 6087329
 GENERAL INFORMATION:
 APPLICANT: ARMITAGE, RICHARD
 APPLICANT: FANSLOW, WILLIAM
 APPLICANT: SPRIGGS, MELANIE
 APPLICANT: SRINIVASAN, SUBHASHINI
 APPLICANT: GIBSON, MARYLOU
 APPLICANT: MORRIS, ARVIA E.
 APPLICANT: MCCREW, JEFFERY
 TITLE OF INVENTION: NOVEL CYTOKINE THAT BINDS CD40
 NUMBER OF SEQUENCES: 26
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: IMMUNEX CORPORATION
 STREET: 51 UNIVERSITY STREET
 CITY: SEATTLE
 STATE: WASHINGTON
 COUNTRY: USA
 ZIP: 98101
 COMPUTER READABLE FORM:
 MEDIUM TYPE: FLOPPY disk
 COMPUTER: Apple Macintosh
 OPERATING SYSTEM: Apple Operating System 7.5.5
 SOFTWARE: MS Word for Apple Power Macintosh, version 6.0.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/088, 913A
 FILING DATE:
 CLASSIFICATION:
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: 08/484, 624
 FILING DATE:
 APPLICATION NUMBER: 08/477, 733
 FILING DATE: June 07, 1995
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: 07/969, 703
 FILING DATE: October 23, 1992
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: 07/805, 723
 FILING DATE: December 5, 1991
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: 07/783, 707
 FILING DATE: October 25, 1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Perkins, Patricia A.
 REGISTRATION NUMBER: 34, 693
 REFERENCE/DOCKET NUMBER: 2802-D
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 2065870430
 TELEFAX: 2065870505
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 261 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-088-913A-12

Query Match 44 6%; Score 110; DB 3; Length 261;
 Best Local Similarity 100.0%; Pred. No. 3. 5e-79;
 Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 257 HRRLDKIEDERNLHEDFVFMKTIQRCNGERSLSLLNCEEIKSOSFGCFVKDMLNKEETK 316
 QY 317 KENSFEMQKGDNQPOIAAHVIBASSKTSVQWAERGYYMMNNVTLENGKQLTVKRQ 376
 Db 107 KENSFEMQKGDNQPOIAAHVIBASSKTSVQWAERGYYMMNNVTLENGKQLTVKRQ 166
 QY 377 GLYYIYAQVFCNRREASSQAPFTASCLSKSPGRERILLRAANTHSSAKPCGQOOSHG 436
 Db 167 GLYYIYAQVFCNSREASSQAPFTASCLSKSPGRERILLRAANTHSSAKPCGQOOSHG 226
 QY 437 GVFELQPGASVNVNTDPSQVSHGTGFTSFGLKL 471

Query Match 44 6%; Score 110; DB 3; Length 261;
 Best Local Similarity 100.0%; Pred. No. 3. 5e-79;
 Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 257 HRRLDKIEDERNLHEDFVFMKTIQRCNGERSLSLLNCEEIKSOSFGCFVKDMLNKEETK 316
 Db 47 HRRLDKIEDERNLHEDFVFMKTIQRCNGERSLSLLNCEEIKSOSFGCFVKDMLNKEETK 106
 QY 317 KENSFEMQKGDNQPOIAAHVIBASSKTSVQWAERGYYMMNNVTLENGKQLTVKRQ 376

05-08-446-922-11

Sequence 11, Application US/08446922

Patent No. 5961974

Patient No. 5716805

GENERAL INFORMATION:

APPLICANT: Spriggs, Melanie

APPLICANT: Srinivasan, Subhashini

TITLE OF INVENTION: Methods of Preparing Soluble, Oligomeric

NUMBER OF INVENTION: Proteins

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Immunex Corporation

STREET: 51 University Street

CITY: Seattle

STATE: WA

COUNTRY: USA

ZIP: 98101

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: Apple Macintosh

OPERATING SYSTEM: Apple Operating System 7.1

SOFTWARE: Microsoft Word for Apple, Version 5.1a

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/446,922

FILING DATE:

CLASSIFICATION: 435

PRIORITY APPLICATION NUMBER: 34,693

REFERENCE/DOCKET NUMBER: USSN 08/107,353

TELECOMMUNICATION INFORMATION:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Perkins, Patricia A

REGISTRATION NUMBER: 34,693

REFERENCE/DOCKET NUMBER: 1003-A

TELEPHONE: (206) 587-0430

TELEFAX: (206) 233-0644

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 273 amino acids

TYPE: amino acid

TOPOLGY: Linear

MOLECULE TYPE: protein

US-08-446-922-11

Query Match

Best Local Similarity 44.0%

Score 1094.5; DB 1; Length 273;

Matches 217; Conservative 8; Mismatches 17; Indels 5; Gaps 2;

QY 226 LQQMELAKGKQLEVAFASHYKQAKLFPDGH RLKDTEDERNLDHEFVFKTQRCNT 284

OY 31 IEDKIEBILSKIYHNETARIKKLI--GERPSSDKIEDERNLDHEFVFKTQRCNT 284

Db 87 GERSLSLNLCEEIKSQEGTVKDMLNKEETKKEENSEFMOKGDNPOTAHVISEASSKT 344

OY 285 GERSLSLNLCEEIKSQEGTVKDMLNKEETKKEENSEFMOKGDNPOTAHVISEASSKT 344

Db 87 GERSLSLNLCEEIKSQEGTVKDMLNKEETKKEENSEFMOKGDNPOTAHVISEASSKT 146

QY 345 TSVLQWAEGGYTMSNNLVTLENGKQLETVKQROGLYYIAQVFCNSREASSQAPFASIC 404

Db 147 TSVLQWAEGGYTMSNNLVTLENGKQLETVKQROGLYYIAQVFCNSREASSQAPFASIC 404

QY 405 LKSPGRFERLILRAANTHSSAKPCGQOSIHLGGVFELOPGASVFNVTDPQVSHGTGFT 464

Db 207 LKSPGRFERLILRAANTHSSAKPCGQOSIHLGGVFELOPGASVFNVTDPQVSHGTGFT 266

QY 465 SFGGLKL 471

Db 267 SFGGLKL 273

Patent No. 5961974

GENERAL INFORMATION:

APPLICANT: ARMITAGE, RICHARD

APPLICANT: FANSLOW, WILLIAM

APPLICANT: SPRIGGS, MELANIE

APPLICANT: SRIINTVASAN, SUBHASHINI

TITLE OF INVENTION: NOVEL CYTOKINE

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:

ADDRESSEE: IMMUNEX CORPORATION

STREET: 51 UNIVERSITY STREET

CITY: SEATTLE

STATE: WASHINGTON

COUNTRY: USA

ZIP: 98101

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: Apple Macintosh

OPERATING SYSTEM: Apple Operating System 7.1

SOFTWARE: Microsoft Word for Apple, version 5.1a

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/249,189

FILING DATE:

CLASSIFICATION: 514

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: 07/805,723

FILING DATE: December 5, 1991

CLASSIFICATION: 514

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: 07/783,707

FILING DATE: October 25, 1991

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Perkins, Patricia A

REGISTRATION NUMBER: 34,693

REFERENCE/DOCKET NUMBER: 2802-C

TELECOMMUNICATION INFORMATION:

TELEPHONE: 206587030

TELEFAX: 2065870505

INFORMATION FOR SEQ ID NO: 21:

SEQUENCE CHARACTERISTICS:

LENGTH: 273 amino acids

TYPE: amino acid

TOPOLGY: Linear

MOLECULE TYPE: protein

US-08-249-189-21

Query Match

Best Local Similarity 44.0%

Score 1094.5; DB 2; Length 273;

Matches 217; Conservative 8; Mismatches 17; Indels 5; Gaps 2;

QY 226 LQQMELAKGKQLEVAFASHYKQAKLFPDGH RLKDTEDERNLDHEFVFKTQRCNT 284

Db 31 IEDKIEBILSKIYHNETARIKKLI--GERPSSDKIEDERNLDHEFVFKTQRCNT 284

QY 285 GERSLSLNLCEEIKSQEGTVKDMLNKEETKKEENSEFMOKGDNPOTAHVISEASSKT 344

Db 87 GERSLSLNLCEEIKSQEGTVKDMLNKEETKKEENSEFMOKGDNPOTAHVISEASSKT 146

QY 345 TSVLQWAEGGYTMSNNLVTLENGKQLETVKQROGLYYIAQVFCNSREASSQAPFASIC 404

Db 147 TSVLQWAEGGYTMSNNLVTLENGKQLETVKQROGLYYIAQVFCNSREASSQAPFASIC 206

QY 405 LKSPGRFERLILRAANTHSSAKPCGQOSIHLGGVFELOPGASVFNVTDPQVSHGTGFT 464

Db 207 LKSPGRFERLILRAANTHSSAKPCGQOSIHLGGVFELOPGASVFNVTDPQVSHGTGFT 266

RESULT 15

US 08-249-189-21

Sequence 21, Application US/08249189

Tue Jul 3 16:26:04 2001

us-09-454-223-2.ra1

Page 9

QY 465 SFGULKL 471
| | | | |
Db 267 SFGULKL 273

Search completed: July 3, 2001, 16:15:31
Job time: 108 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2001, 16:14:03 ; Search time 20.96 seconds

{without alignments} (1711.747 Million cell updates/sec)

Title: US-09-454-223-2
Perfect score: 2487

Sequence: 1 MPPFLSMLVLLVOPGLGNLGA. TDPSSQVSHGTGFTSFGLLKL 471

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241
Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68.*
1: p1x1;*
2: p1x2;*
3: p1x3;*
4: p1x4;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No. Score Query Match Length DB ID Description
1 1286 51.7 374 1 A42046 surfactant protein D - rat
2 1110 44.6 261 2 I53476 Cb40 ligand - huma
3 1053 42.3 375 1 A45225 pulmonary surfacta
4 974 39.2 261 2 S53090 Cb40 ligand - bovi
5 961.5 38.7 369 2 S33603 surfactant protein
6 819 35.3 371 2 I45878 conglutinin - bovi
7 878 35.3 371 1 JN0450 conglutinin precur
8 827 33.3 260 2 S21738 Cb40 ligand - mous
9 565.5 22.7 301 2 A53570 collectin-43 - bov
10 488.5 19.6 888 2 S28791 collagen alpha 1(X
11 484.5 19.5 1049 1 CGB075 collagen alpha 1(T
12 481 19.3 1042 1 CGCH13 collagen alpha 1(T
13 1806 19.3 1 CGHUI1 collagen alpha 1(X
14 479.5 19.3 1838 1 CGHUIV collagen alpha 1(V
15 479.5 19.3 1843 2 S18803 collagen alpha 1(V
16 477.5 19.2 1464 2 S59836 collagen alpha 1(T
17 477 19.2 671 1 CGRT13 collagen alpha 1(T
18 476.5 19.2 1453 2 S21262 collagen alpha 1(T
19 476 19.1 1532 2 A61262 collagen alpha 1(X
20 475 19.1 1433 2 A46053 collagen alpha 1(V
21 474.5 19.1 1464 1 CGHUI1 collagen alpha 1(V
22 474.5 19.1 1466 1 CGHUI7 collagen alpha 1(T
23 473.5 19.0 1487 1 CGHUI6C collagen alpha 1(T
24 473.5 19.0 1546 1 CGHUI2 collagen alpha 2(X
25 473 19.0 1492 2 A40333 collagen alpha 1(T
26 472.5 19.0 779 1 CGB013 collagen alpha 1(T
27 472.5 19.0 1763 2 S16366 collagen alpha 2(T
28 472 19.0 1418 2 A45467 collagen alpha 1(T
29 471 18.9 1486 1 B40333 collagen alpha 1(T

ALIGNMENTS

RESULT 1
A42046 surfactant protein D - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 10-Sep-1999 #sequence_revision 10-sep-1999
C;Accession: A42046
R;Shimizu, H.; Fisher, J.H.; Papst, P.; Benson, B.; Lau, K.; Mason, R.J.; Voelker, D.
J. Biol. Chem. 267: 1853-1857, 1992
A;Title: Primary structure of rat pulmonary surfactant protein D. cDNA and deduced am
A;Reference number: A42046; MUID:92112913
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-374 <SH1>
A;Cross references: GB:W81211; MUID:9207035; PIDN:AAA42170.1; PID:9207035
A;Experimental source: Lung
A;Note: sequence extracted from NCBI backbone (NCBIN:76027, NCBIPI:6031)
C;Superfamily: pulmonary surfactant protein D; C-type lectin homology
F;253-372/Domain: C-type lectin homology <LCH>

Query Match 51.7%; Score 1286; DB 1; Length 374;
Best Local Similarity 85.7%; Pred. No. 1.5e-72;
Matches 245; Conservative 7; Mismatches 34; Indels 0; Gaps 0;

QY 1 MPPFLSMLVLLVOPGLGNLGA. TDPSSQVSHGTGFTSFGLLKL 471
Db 1 MPPFLSMLVLLVOPGLGNLGA. TDPSSQVSHGTGFTSFGLLKL 471

QY 61 EKGDPGLCPGPMGLSGLQQPTGPVYKGPKGNSAGEPGPKGERGLSGPGLPQGPGPACKEG 120
Db 61 EKGDPGLCPGPMGLSGLQQPTGPVYKGPKGNSAGEPGPKGERGLSGPGLPQGPGPACKEG 120

QY 121 PSGKGNIGQPGKPGPKGEAGPKGEVAGPAGMKGSTGKSTGKGERGAPGVYQAGPNAG 180
Db 121 PSGKGNIGQPGKPGPKGEAGPKGEVAGPAGMKGSTGKSTGKGERGAPGVYQAGPNAG 180

QY 181 AAGPAGPAGQAGPSRGPGLKGDRGFGDRGKGERSGLPDSALRQOMEALKGKQRR 240
Db 181 AAGPAGPAGQAGPSRGPGLKGDRGFGDRGKGERSGLPDSALRQOMEALKGKQRR 240

QY 241 EVAFSHYKKAALPDGPHRLDKFEDERNMHEDYFMKLTORCNGE 286
Db 241 EAAPFSRYKKAALPDGQSVGDKFRAANNEEPFEDAKEMCRQAGQ 286

RESULT 2

15476
CDB01 ligand - human
N;Alternate names: glycoprotein 39; hCD40-L protein; T-cell antigen gp39; TRAP protein
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 08-Oct-1999

C;Accession: 153476; S28017; JH0793; S26694; S28852; S25684; S30593		Db 227 GVFELOPGASVFVNTPSQVSHGFTSFGLKL 261
R;Gauthat, J.	FEBS Lett. 315, 259-266, 1992	
A;Title: Human CD40-ligand: Molecular cloning, cellular distribution and regulation of e	A;Status: preliminary; translated from GB/EMBL/DDJB	
A;Accession: 153476; MUID:93138085	A45225	
A;Molecule type: mRNA	A45225	
A;Residues: 1-261 <RES>	N;Alternative names: SP-D	
A;Cross-references: GB:L07414; NID:918023; PIDN:AAA35662.1; PID:9180124	C;Species: Homo sapiens (man)	
A;Title: The human T cell antigen gp39, a member of the TNF gene family, is a ligand for	C;Date: 16-Apr-1999 "sequence_revision 16-Apr-1999 #text-change 22-Jun-1999	
A;Accession number: S28017; MUID:93049181	C;Accession: A45225; S23434; S44555; S44620; S18382; A50776	
A;Accession: S28017	R;Crouch, E.; Rust, K.; Vandeplas, R.; Denis Keller, H.; Grosso, L.	
A;Molecule type: mRNA	J. Biol. Chem. 268, 2976-2983, 1993	
A;Residues: 1-261 <RQL>	A;Title: Genomic organization of human surfactant protein D (SP-D). SP-D is encoded o	
A;Cross-references: EMBL:Z15017; NID:930493; PIDN:CAA7837.1; PID:938084	A;Reference number: A45225; MUID:93155122	
R;Spriggs, M.K.; Armitage, R.J.; Schroepke, L.; Clifford, K.N.; Macduff, B.M.; Sato, T.	A;Accession: A45225	
J. Exp. Med. 176, 1543-1550, 1992	A;Molecule type: DNA	
A;Title: Recombinant human CD40 ligand stimulates B cell proliferation and immunoglobuli	A;Residues: 1-375 <CRQ>	
A;Reference number: JH0793; MUID:93094757	A;Cross-references: GB:L05483; GB:L05484; NID:9292505; PIDN:AAB59450.1; PI	
A;Accession: JH0793	A;Experimental source: placenta	
A;Molecule type: mRNA	A;Note: sequence extracted from NCBI backbone (NCBIP:124316)	
A;Residues: 1-261 <GRA>	R;Lu, J.; Willis, A.C.; Reid, K.B.M.	
A;Cross-references: EMBL:X69550; NID:937269; PIDN:CAA48554.1; PID:937270	Biochem. J. 284, 795-802, 1992	
A;Title: Human CD40 ligand: molecular cloning, cellular distribution and regulation of e	A;Accession: S28852	
A;Reference number: S28852; MUID:93138085	A;Title: Purification, characterization and cDNA cloning of human lung surfactant pro	
A;Molecule type: mRNA	A;Reference number: S23434; MUID:92322003	
A;Residues: 1-261 <GRA>	A;Accession: S23434	
A;Cross-references: EMBL:L07414; NID:9180123; PIDN:AAA35662.1; PID:9180124	A;Molecule type: mRNA	
A;Title: the sequence from Fig. 3 is inconsistent with that from Fig. 2 in having 6-Gln	A;Residues: 1-30, 'T', 32-121, 'P', 123-179, 'A', 181-375 <LUJ1>	
A;Reference number: S28852; MUID:93138085	A;Cross-references: EMBL:X65018; NID:934766; PIDN:CAA46152.1; PID:934767	
A;Accession: S28852	A;Experimental source: lung	
A;Molecule type: mRNA	A;Accession: S44555	
A;Residues: 1-261 <GRA>	A;Molecule type: protein	
A;Cross-references: EMBL:L07414; NID:9180123; PIDN:AAA35662.1; PID:9180124	A;Residues: 214-234, 'X', 236, 'XX', 239-241 <LUJ2>	
A;Title: Human CD40 ligand: molecular cloning, cellular distribution and regulation of e	R;Hoipe, R.J.; Barlow, P.N.; Reid, K.B.M.	
A;Reference number: S28852; MUID:93138085	FEBS Lett. 315, 259-266, 1992	
A;Accession: S28852	A;Title: A parallel three stranded alpha-helical bundle at the nucleation site of col	
A;Molecule type: mRNA	A;Residues: 191-195, 1994	
A;Residues: 1-261 <GRA>	R;Rust, K.; Gross, L.; Zhang, V.; Chang, D.; Persson, A.; Longmore, W.; Cai, G.Z.; C	
A;Cross-references: EMBL:L07414; NID:9180123; PIDN:AAA35662.1; PID:9180124	A;Title: Human surfactant protein D: SP-D contains a C-type lectin carbohydrate recog	
A;Title: Accumulation of surfactant protein D in human pulmonary alveolar proteinosis	A;Reference number: S18382; MUID:91378578	
A;Accession: S18382	A;Accession: S4420	
A;Molecule type: mRNA	A;Residues: 202-257 <HOP>	
A;Status: preliminary	R;Rust, K.; Gross, L.; Zhang, V.; Chang, D.; Persson, A.; Longmore, W.; Cai, G.Z.; C	
A;Molecule type: mRNA	A;Title: Human surfactant protein D: SP-D contains a C-type lectin carbohydrate recog	
A;Residues: 'P', 60-205, 'P', 207-374, 'HF' <RUS>	A;Reference number: S18382; MUID:91378578	
A;Cross-references: GB:L05455; NID:9222505	A;Accession: S18382	
A;Note: corrections to this sequence are reported in reference A56776	A;Title: Human surfactant protein D: SP-D contains a C-type lectin carbohydrate recog	
R;Crouch, E.; Persson, A.; Chang, D.	A;Residues: 'P', 60-205, 'P', 207-374, 'HF' <RUS>	
A;Title: Accumulation of surfactant protein D in human pulmonary alveolar proteinosis	A;Reference number: S18382; MUID:91378578	
A;Accession: A56776	A;Accession: S18382	
A;Status: preliminary	A;Title: Human surfactant protein D: SP-D contains a C-type lectin carbohydrate recog	
A;Molecule type: protein	A;Accession: S18382	
A;Residues: 'P', 60-62, 'E', 64-72, 223-227, 'X', 229-239, 'P', 241-245, 'X', 247-256, 'X'	A;Title: Human surfactant protein D: SP-D contains a C-type lectin carbohydrate recog	
A;Cross-references: PIDN:AA25037.1; PID:2636973; PIDN:AA25038.1; PID:9263974	A;Accession: S18382	
A;Experimental source: bronchoalveolar lavage	A;Title: Human surfactant protein D: SP-D contains a C-type lectin carbohydrate recog	
A;Note: sequence extracted from NCBI backbone (NCBIP:123024, NCBIP:123023); sequence	A;Accession: S18382	
C;Comment: Pulmonary surfactant is a complex of phospholipids and proteins that lower	A;Title: Human surfactant protein D: SP-D contains a C-type lectin carbohydrate recog	
C;Genetics:	A;Accession: S18382	
A;Cross-references: SFRP4; SP-D	A;Title: Human surfactant protein D: SP-D contains a C-type lectin carbohydrate recog	
A;Map Position: 10q22.2-10q23.1	A;Accession: S18382	
C;Superfamily: Pulmonary surfactant protein D, C-type lectin homology	A;Title: Human surfactant protein D: SP-D contains a C-type lectin carbohydrate recog	
C;Keywords: blocked amino end; calcium; glycoprotein; hydroxylysine; hydroxyproline;	A;Accession: S18382	
F;1-20;Domain: signal sequence #status predicted <SIG>	A;Title: Human surfactant protein D: SP-D contains a C-type lectin carbohydrate recog	
F;21-375/Product: pulmonary surfactant protein D #status predicted <MAT>	A;Accession: S18382	
F;46-222/Domain: non-collagenous #status predicted <COL>	A;Title: Human surfactant protein D: SP-D contains a C-type lectin carbohydrate recog	

A; Reference number: S35044; MUID:93358905
 A; Accession: S35044
 A; Molecule type: protein
 A; Residues: 75-'X', 'Y', '88-99, 'X', '91, 'T' <LUA>
 A; Experimental source: lung
 R; Young, N.M.; Leon, M.A.
 Blochem. Biophys. Res. Commun. 143, 645-651, 1987
 A; Title: The carbohydrate specificity of conglutinin and its homology to proteins in
 A; Reference number: A29416; MUID:87184551
 A; Contents: annotation
 R; Malhotra, R.; Laursen, S.B.; Willis, A.C.; Sim, R.B.
 Blochem. J. 239, 15-19, 1993
 A; Title: Research Communication. Localization of the receptor-binding site in the colo-
 A; Reference number: S34054; MUID:93319501
 A; Contents: annotation
 R; Liou, L.S.; Sastry, R.; Hartshorn, K.L.; Lee, Y.M.; Okarma, T.B.; Tauber, A.I.; Sas-
 J. Immunol. 153, 173-180, 1994
 A; Title: Bovine conglutinin gene exon structure reveals its evolutionary relationship
 A; Reference number: I46010; MUID:94267222
 A; Accession: I46010
 A; Status: preliminary; translated from GB/EMBL/DDBJ
 A; Molecule type: DNA
 A; Residues: 1-371 <LIO>
 A; Cross-references: EMBL:U05860; NID:9507183; PIDN:AB60624.1; PIB:9514256
 C; Comment: This protein mediates the agglutination of erythrocytes with antibody and
 C; Genetics: This protein is a Ca2+-dependent serum lectin specific for N-acetylglucosa-
 A; Gene: CGN1
 A; Introns: 67/1; 106/1; 142/1; 178/1; 217/1; 245/1
 C; Superfamily: pulmonary surfactant protein D; C-type lectin homology
 C; Keywords: calcium binding; glycoprotein; hydroxylysine; hydroxyproline
 F; 1-20/Domain: signal sequence #status predicted <SIG>
 F; 21-371/Product: conglutinin #status predicted <MAT>
 F; 46-214/Region: collagen-like
 F; 75-371/Product: conglutinin-N #status predicted <MAT>
 F; 248-369/Domain: C-type lectin homology <LCH>
 F; 63, 87, 99, 135, 141, 150, 162, 198, 210/Binding site: carbohydrate (Lys) (covalent) #status exper-
 F; 63, 87, 99, 135, 141, 150, 162, 198, 210/Binding site: 5-hydroxylysine (Lys) #status exper-
 F; 8, 96, 108, 111, 129, 132, 147, 153, 171, 175/Modified site: 4-hydroxyproline (Pro) #status
 F; 337/Binding site: carbohydrate (Asn) (covalent) #status experimental
 A; Cross-references: DDBJ:D14085; NID:9285643; PIDN:BAA03170.1; PIB:9285644
 A; Experimental source: liver
 R; Kawasaki, N.; Itoh, N.; Karasaki, T.; Wakamiya, N.
 Blochem. Biophys. Res. Commun. 198, 597-604, 1994
 A; Title: Gene organization and 5'-flanking region sequence of conglutinin: A C-type mammalian conglutinin gene
 A; Reference number: JCB2396; MUID:94128104
 A; Accession: JCB2396
 A; Molecule type: mRNA
 A; Residues: 1-371 <KA2>
 A; Note: The authors translated the codon GAT for residues 250 and 270 as Glu
 R; Lu, J.; Laursen, S.B.; Thiel, S.; Jensenius, J.C.; Reid, K.B.M.
 Blochem. J. 202, 157-162, 1993
 A; Title: the cDNA cloning of conglutinin and identification of liver as a primary site of
 A; Residues: 1-371 <KA2>
 A; Accession: S33235; MUID:93277452
 A; Molecule type: mRNA
 A; Residues: 1-172, 'H', 174-217, 'A', 219-271, 'V', 273-371 <LUJ>
 A; Cross-references: EMBL:X17174; NID:9395267; PIDN:CAA0665.1; PIB:9395268
 A; Experimental source: liver
 R; Lee, Y.M.; Leiby, K.R.; Allar, J.; Paris, K.; Lerch, B.; Okarma, T.B.
 J. Biol. Chem. 266, 2715-2723, 1991
 A; Title: Primary structure of bovine conglutinin, a member of the C-type animal lectin family
 A; Reference number: A23740; MUID:91131556
 A; Accession: A23740
 A; Molecule type: protein
 A; Residues: 21-219, 'S', 211-371 <LEE>
 R; Kawasaki, N.; Yokota, Y.; Kawasaki, T.
 Arch. Biochem. Biophys. 305, 533-540, 1993
 A; Title: Differentiation of conglutination activity and sugar-binding activity of conglutinin
 A; Reference number: S36879; MUID:93384312
 A; Accession: S36879
 A; Molecule type: protein
 A; Residues: 21-54, 75-86, 'X', 88-89, 'X', 91, 'X', 93-94; 208-209, 'X', 211-227 <KAW>
 R; Lu, J.; Wiedemann, H.; Holmskov, U.; Thiel, S.; Timpl, R.; Reid, K.B.M.
 Eur. J. Biochem. 215, 795-799, 1993
 A; Title: Structural similarity between lung surfactant protein D and conglutinin

RESULT 8

S21738

CD40 ligand - mouse

C;Species: Mus musculus (house mouse)

C;Accession: S21738 #sequence_revision 31-Dec-1993 #text_change 05-Nov-1999

C;Accession: S21738 R.Amitage, R.J.; Fanslow, W.C.; Strockbine, L.; Sato, T.A.; Clifford, K.N.; Macduff, B.

; Cosman, D.; Spring, M.K.

Nature 357, 80-82, 1992

A;Title: Molecular and biological characterization of a murine ligand for CD40.

A;Reference number: S21738; MUID:92244364

A;Molecule type: mRNA

A;Residues: 1-260 <RMS>

A;Cross-references: EMBL: X65453; NID:950351; PIDN:CAA46448.1; PID:950352

A;Keywords: glycoprotein; transmembrane #status predicted <TM>

A;F:7-260/Domain: transmembrane #status predicted <EX>

A;F:239/Binding site: carbohydrate (Asn) (covalent) #status predicted

A;Note: sequence extracted from NCBI backbone (NCBIP:131234)

Query Match 22.7%; Score 565.5; DB 2; Length 301; Best Local Similarity 35.8%; Pred. No. 4.1e-28; Matches 127; Conservative 42; Mismatches 81; Indels 105; Gaps 8; Nature 357, 80-82, 1992

Query 21 EMKSLQSNSVNPNTLVMCSPTENGTPDRGDRGREGRRGEKQDPGPGMGLSGQPT 80

Db 2 ENVYSEERLTIDCCTLVACPPDSLRLRGDRGREGRRGEKQDP----- 46

Db 47 ----- 46

Db 81 GPGPKGNGSAGEPGPKGERGLSGPGLPGPGPAGKPGKGEA 140

Db 47 ----- 46

Db 80 GPGVGVAGPMGS--PGPAGLKGKGERGAP-----GPGGAIGPOCRSGAMGPP 124

Db 201 GLKGDRGVPGDGTGKGEGLGPDASALRQOMEALKGKQLRLEVAFPSHQKAALPDKHRL 260

Db 125 GLKGDRGVPGDGTGKGEGLGPDASALRQOMEALKGKQLRLEVAFPSHQKAALPDKHRL 184

Db 261 DKTEDERNHEDFEMKTIORCWGRSLLNCEEEKTSQF-----EGF 304

Db 185 EKI-----RKTAGWVSKYSDAEQL-----CREAKGKLSAPRSSAENEAVTOLVRAK 230

Db 305 VDKDMLNEETKENSEFMQK-----DOPQIAHWISEAS 341

Db 231 NKHAYLSDMDISKEGKFPTGGSDYSNWAPGEPNNRAKDEGPNCELEYSDCN 285

RESULT 9

A5570

A5570 N;Alternate names: lectin CL-43

C;Species: Bos primigenius taurus (cattle)

C;Accession: 12-Apr-1995 #sequence_revision 23-Feb-1996 #text_change 17-Mar-1999

C;Accession: A5350; R66689

Ridm, B.L.; Willis, A.C.; Reid, K.B.M.; Lu, J.; Laursen, S.B.; Jensenius, J.C.; Holmsk

J. Biol. Chem. 269, 11620-11824, 1994

A;Title: Primary structure of bovine collectin-43 (CL-43). Comparison with conglutinin

A;Reference number: A53570; MUID:94216283

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-888 <NAH>

A;Cross-references: EMBL: M88593; NID:9211619; PIDN:AAA46707.1; PID:9211620

C;Superfamily: collagen alpha 1(V) chain; fibrillar collagen carboxyl-terminal homology <FC>

C;Superfamily: collagen alpha 1(V) chain; fibrillar collagen carboxyl-terminal homology <FC>

F:665-887/Domain: fibrillar collagen carboxyl-terminal homology

Query Match 19.6%; Score 488.5; DB 2; Length 888; Best Local Similarity 47.7%; Pred. No. 8.3e-23; Matches 102; Conservative 17; Mismatches 62; Indels 33; Gaps 4; Nature 357, 80-82, 1992

Query 41 PTENGLPGDGRGREGREGEKGDPGPGCPMGSGLQGPTGPVGPK-----GE 88

Db 399 PGPGPGERAGPQDGPGVGGKGEDPGQCPGPGCGEAPGPGPKRPPGATGAEGROE 458

Db 89 NGSGAGE-----GPKBGRGLSGPPGPGPAGK-----GPGKOGNTGP 130

Db 459 KGAKGEPGEGAGPCKTGPVGPQGAGPKCPPEGLGPGVNGEGLPAPGDPGCPPGHLGP 518

Query 131 OGPGPKGERAGPQDGPGVGGKGEDPGQCPGPGCGEAPGPGPKRPPGATGAEGROE 187

Db 519 PGJPGLKGPGSKGEKGHGHLGPGPQGKCPPEGLGPGVNGEGLPAPGDPGCPPGHLGP 578

Query 188 AGQGAPGSRGPGQKLGPGVGPQGAGPKCPPEGLGPGVNGEGLPAPGDPGCPPGHLGP 221

Db 579 LGPGPPEUPGPQGPKGSQGSSPAGQGDGSP 612

C;Superfamily: pulmonary surfactant protein D; C-type lectin homology

C;Keywords: lectin C-type lectin homology <LCH>

F:177-299/Domain: C-type lectin homology <LCH>

C;Species: Primigenius taurinus (cattle)
 C;Date: 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change 07-May-1999
 C;Accession: A02862; A38001; A38002; A38003; A38004; A38005; S71946
 R;Fietzek, P.P.; Altmann, H.; Rauhberg, J.; Henkell, W.; Wachter, E.; Kuehn, K.
 Hoppe-Seyler's Z. Physiol. Chem. 360, 809-820, 1979
 A;Title: The covalent structure of calf skin type III collagen. I. The amino acid sequence
 A;Reference number: A02862; MUID:80026026
 A;Molecule type: protein
 A;Residues: 1-242 <PDB>
 R;Dewes, H.; Fietzek, P.P.; Kuehn, K.
 Hoppe-Seyler's Z. Physiol. Chem. 360, 821-832, 1979
 A;Title: The covalent structure of calf skin type III collagen. II. The amino acid sequence
 A;Reference number: A38001; MUID:80026027
 A;Molecule type: protein
 A;Residues: 243-422 <DBW1>
 R;Bentz, H.; Fietzek, P.P.; Kuehn, K.
 Hoppe-Seyler's Z. Physiol. Chem. 360, 833-840, 1979
 A;Title: The covalent structure of calf skin type III collagen. III. The amino acid sequence
 A;Reference number: A38002; MUID:80026028
 A;Molecule type: protein
 A;Accession: A38002
 A;Residues: 423-571 <PDB>
 R;Lang, H.; Glanville, R.W.; Fietzek, P.P.; Kuehn, K.
 A;Title: The covalent structure of calf skin type III collagen. IV. The amino acid sequence
 A;Reference number: A38003; MUID:80026029
 A;Accession: A38003
 A;Residues: 572-808 <PDB>
 R;Dewes, H.; Fietzek, P.P.; Kuehn, K.
 Hoppe-Seyler's Z. Physiol. Chem. 360, 851-860, 1979
 A;Title: The covalent structure of calf skin type III collagen. V. The amino acid sequence
 A;Reference number: A38004; MUID:80026030
 A;Accession: A38004
 A;Molecule type: protein
 A;Residues: 809-947 <PDB>
 R;Altmann, H.; Fietzek, P.P.; Kuehn, K.
 Hoppe-Seyler's Z. Physiol. Chem. 360, 861-868, 1979
 A;Title: The covalent structure of calf skin type III collagen. VI. The amino acid sequence
 A;Reference number: A38005; MUID:80026031
 A;Accession: A38005
 A;Molecule type: protein
 A;Residues: 948-149 <PDB>
 A;Experimental source: skin
 A;Henzel, W.
 A;Lochem, J. 318, 497-503, 1996
 A;Title: Cross-link analysis of the C-telopeptide domain from type III collagen.
 A;Reference number: S1946; MUID:96404897
 A;Accession: S71946
 A;Molecule type: protein
 A;Residues: 87-105; 107-1029; 1037-1049 <HRN>
 A;Comment: Prolines at the third position of the tripeptide repeating unit (G-X-Y) are important for the covalent structure of type III collagen molecule.
 A;Comment: The type III collagen molecule is a trimer of identical chains, linked to each other by noncovalent interactions.
 A;Keywords: collagen alpha 1(III) chain; fibrillar collagen; carboxyl-terminal homology; 1-1099/product; collagen alpha 1(III) chain #status experimental <CAB>
 A;1-14/Region: amino-terminal nonhelical telopeptide
 A;15-1040/Region: helical
 A;587-589/Region: cell attachment (R-G-D) motif
 A;752-754/Region: cell attachment (R-G-D) motif
 A;875-877/Region: cell attachment (R-G-D) motif
 A;878-880/Region: cell attachment (R-G-D) motif
 A;935-937/Region: cell attachment (R-G-D) motif
 A;951-1049/Region: carbonyl-terminal nonhelical telopeptide
 A;95, 107, 119, 938, 950/Modified site: 5-hydroxylysine (Lys) #status predicted
 A;107, 950/Modified site: allysine (Lys) #status predicted
 A;107/Binding site: carbohydrate (Lys) (covalent) #status experimental

RESULT 13

Db 487 AGPPGEAKRPGEOGPGNAGAPGPGAGARGERGPFGERGVQGPQGPQPRGANGAPGNDGA 546

QY 215 KGSSGLP 221

Db 547 KGAGAP 553

CGHUV

N:Alternate names: procollagen alpha₁(XI) chain

C:Species: Homo sapiens (man)

C;Date: 31-Mar-1990 #sequence_revision 03-Oct-1995 #text_change 08-May-1998

C:Accession: A35239; A31795

R;Roshko, H.; Ramirez, F.

J; Biol. Chem. 265, 6423-6426, 1990

A;Title: Pro-alpha1(XI) collagen. Structure of the amino-terminal propeptide and expression

A;Reference number: A35239; MUID:90202924

A;Molecule type: mRNA

A;Residues: 1-558 <YOS>

A;Cross-references: GB:J05407

R;Bernard, M.; Yoshikiri, H.; Rodriguez, E.; van der Rest, M.; Kimura, T.; Niromiya, Y.; J. Biol. Chem. 263, 17159-17166, 1988

A;Title: Cloning and sequencing of pro-alpha1(XI) collagen cDNA demonstrates that type XI collagen is encoded by a single gene

A;Accession: A31795

A;Molecule type: DNA; mRNA

A;Residues: 538-1806 <8ER>

A;Cross-references: GB:J04177

A;Title: parts of this sequence were determined by protein sequencing

C;Comment: Prolines and lysines at the third position of the tripeptide repeating unit are O-glycosylated and subsequently O-glycosylated.

C;Genetics

A;Gene: GDB:COL11A1; COL11

A;Cross-references: GDB:120595; OMIM:120280

A;Map position: 1p21-1p21

A;Intron: 561/3; 579/3; 597/3; 615/3; 633/3; 648/3; 666/3; 681/3

A;Note: the list of introns is incomplete

C;Complex: type XI collagen may be a heterotrimer of two alpha 1(XI) chains and one alpha 3(XII) chain (see PIR:CHHUC), initially linked by disulfide bonds among their carboxyl-terminal with desmosine cross-links made from lysine and allysine residues

C;Function:

A;Description: structural component of extracellular fibrous polymer associated with cell

A;Note: may play a role in controlling the lateral growth of collagen II fibrils

C;Superfamily: collagen alpha 1(V) chain; fibrillar collagen carboxyl-terminal homology

C;Keywords: coiled coil; extracellular matrix; glycoprotein; hydroxylsine; hydroxyproline; desmosine; cross-links; collagen

F;1-36/Domain: signal sequence #status predicted <SIC>

F;26-260/Domain: PARP-like #status predicted <PARP>

F;37-51/Domain: amino-terminal propeptide #status predicted <PRO>

F;51-155/Protein: collagen alpha 1(XI) chain #status predicted <MAT>

F;51-52/Region: amino-terminal nonhelical telopeptide

F;128-152/Region: helical

F;156-1565/Region: carboxyl-terminal nonhelical telopeptide

F;1566-1806/Domain: carboxyl-terminal propeptide #status predicted <CTP>

F;1583-1805/Domain: fibrillar collagen carboxyl-terminal homology <FCC>

F;161-243-182-236/Disulfide bonds: #status predicted

F;161-243-182-236/Modified site: allysine (Lys) #status predicted

F;612-1452/Modified site: 5-hydroxylysine (Lys) #status predicted

F;612-1452/Binding site: carbohydrate (Lys) (covalent) #status predicted

• Query Match, 19.3%; Score 480; DB 1; Length 1806;

Best Local Similarity 45.5%; Pred. No. 6.4e-22;

Matches 100; Conservative 18; Mismatches 60; Indels 42; Gaps 4;

QY 41 PTENGLPGR--DARDGREGPRERKGDPDGLPGMGLQGPPGP----- 82

Db 1314 PGDPGPPGGLGPGQPGVQGKGDKDQCPGPGPSGEAGPQCPGKRGPPCAAGER 1373

QY 83-VGPKENGNSAGEPGPKGERGLUSGPPGLPGTPCPAGKEPGSGK 124

RESULT 14

Db 1374 OGEGKAKGEAGABGPPGKTKTPVPGPQGPAGKPGPEGLGPGPVGPEOGLPQGAQGDSPRG 1433

QY 125 QGNITGPOGKPGPKGEAGKPGKEVAPGPMQGSGTAKGSTGKGSGTGPGERGAPGVQAP--GUAGA 181

Db 1434 -- MGPPSGLPGKDPGKGKEKGHPLGLIGLPGGEQEGKGRGLPFGQSGKAGOGI 1490

QY 182 AGPAGPAGPQGAGSGSRGPPGLKDRGKKGESGLP 221

Db 1491 PGPAKPLGPGLGPGLGPQGPKGNKGSTGPGAGQKGDSGLP 1530

CGHUV

N:Alternate names: procollagen alpha 1(V) chain

C:Species: Homo sapiens (man)

C;Date: 22-Nov-1993 #sequence_revision 03-Oct-1995 #text_change 16-Jun-2000

C;Accession: S18802; S16024; A61142; S11303; S03978; S43642; S58665

R;Greenspan, D.S.; Cheng, W.; Hoffman, G.G.

J; Biol. Chem. 266, 24727-24733, 1991

A;Title: The pro-alpha1(V) collagen chain. Complete primary structure, distribution and properties

A;Reference number: S18802; MUID:92105142

A;Accession: S18802

A;Molecule type: mRNA

A;Residues: 1-1838 <GRE>

A;Cross-references: GB:W7679; NID:9189519; PID:AAA59993.1; PID:9189520

R;Takahara, K.; Sato, Y.; Ozawa, K.; Okamoto, N.; Noda, A.; Yaoi, Y.; Kato, I.

J; Biol. Chem. 266, 13124-13129, 1991

A;Title: Complete primary structure of human collagen alpha-1(V) chain

A;Accession: S16024; MUID:91302336

A;Molecule type: mRNA

A;Residues: 1-81, 'Q', 'A', '391-676', 'K', '678-1294', 'P', '1297', 'R', '1300-1553', 'R', '15

A;Cross-references: GB:D90229; NID:9219509; PID:BA14323.1; PID:9219510

A;Title: parts of this sequence were determined by protein sequencing

R;Yaoi, Y.; Hashimoto, K.; Takahara, K.; Kato, I.

Exp. Cell Res. 194, 180-185, 1991

A;Title: Insulin binds to type V collagen with retention of mitogenic activity

A;Reference number: A61142; MUID:91224163

A;Accession: A61142

A;Molecule type: protein

A;Residues: 823-824, 'X', '826-842 <YAO>

A;Note: the residue designated 'X' is probably glycosylated hydroxylsine; this cyano

C;Function:

A;Description: structural component of extracellular fibrous polymer associated with cell

A;Note: may play a role in controlling the lateral growth of collagen II fibrils

C;Complex: type XI collagen may be a heterotrimer of two alpha 1(XI) chains and one alpha 3(XII) chain (see PIR:CHHUC), initially linked by disulfide bonds among their carboxyl-terminal with desmosine cross-links made from lysine and allysine residues

C;Function:

A;Title: parts of this sequence were determined by protein sequencing

R;Yaoi, Y.; Hashimoto, K.; Kato, I.

Arch. Biochem. Biophys. 271, 120-129, 1989

A;Title: Primary structure of the heparin-binding site of type V collagen

A;Reference number: S11303; MUID:90366601

A;Accession: S11303

A;Molecule type: protein

A;Residues: 823-824, 'X', '826-848', 'I', '850-851', 'P', '853', 'PR', '856-893', 'D', '895-932', 'X', '934-935', 'R', '936-940', 'P', '941-945', 'K', '946-950', 'R', '951-955', 'P', '956-960', 'K', '961-965', 'R', '966-970', 'P', '971-975', 'K', '976-980', 'R', '981-985', 'P', '986-990', 'K', '991-995', 'R', '996-998', 'P', '999-1000', 'K', '1001-1002', 'R', '1003-1004', 'P', '1005-1006', 'K', '1007-1008', 'R', '1009-1010', 'P', '1011-1012', 'K', '1013-1014', 'R', '1015-1016', 'P', '1017-1018', 'K', '1019-1020', 'R', '1021-1022', 'P', '1023-1024', 'K', '1025-1026', 'R', '1027-1028', 'P', '1029-1030', 'K', '1031-1032', 'R', '1033-1034', 'P', '1035-1036', 'K', '1037-1038', 'R', '1039-1040', 'P', '1041-1042', 'K', '1043-1044', 'R', '1045-1046', 'P', '1047-1048', 'K', '1049-1050', 'R', '1051-1052', 'P', '1053-1054', 'K', '1055-1056', 'R', '1057-1058', 'P', '1059-1060', 'K', '1061-1062', 'R', '1063-1064', 'P', '1065-1066', 'K', '1067-1068', 'R', '1069-1070', 'P', '1071-1072', 'K', '1073-1074', 'R', '1075-1076', 'P', '1077-1078', 'K', '1079-1080', 'R', '1081-1082', 'P', '1083-1084', 'K', '1085-1086', 'R', '1087-1088', 'P', '1089-1090', 'K', '1091-1092', 'R', '1093-1094', 'P', '1095-1096', 'K', '1097-1098', 'R', '1099-1100', 'P', '1101-1102', 'K', '1103-1104', 'R', '1105-1106', 'P', '1107-1108', 'K', '1109-1110', 'R', '1111-1112', 'P', '1113-1114', 'K', '1115-1116', 'R', '1117-1118', 'P', '1119-1120', 'K', '1121-1122', 'R', '1123-1124', 'P', '1125-1126', 'K', '1127-1128', 'R', '1129-1130', 'P', '1131-1132', 'K', '1133-1134', 'R', '1135-1136', 'P', '1137-1138', 'K', '1139-1140', 'R', '1141-1142', 'P', '1143-1144', 'K', '1145-1146', 'R', '1147-1148', 'P', '1149-1150', 'K', '1151-1152', 'R', '1153-1154', 'P', '1155-1156', 'K', '1157-1158', 'R', '1159-1160', 'P', '1161-1162', 'K', '1163-1164', 'R', '1165-1166', 'P', '1167-1168', 'K', '1169-1170', 'R', '1171-1172', 'P', '1173-1174', 'K', '1175-1176', 'R', '1177-1178', 'P', '1179-1180', 'K', '1181-1182', 'R', '1183-1184', 'P', '1185-1186', 'K', '1187-1188', 'R', '1189-1190', 'P', '1191-1192', 'K', '1193-1194', 'R', '1195-1196', 'P', '1197-1198', 'K', '1199-1200', 'R', '1201-1202', 'P', '1203-1204', 'K', '1205-1206', 'R', '1207-1208', 'P', '1209-1210', 'K', '1211-1212', 'R', '1213-1214', 'P', '1215-1216', 'K', '1217-1218', 'R', '1219-1220', 'P', '1221-1222', 'K', '1223-1224', 'R', '1225-1226', 'P', '1227-1228', 'K', '1229-1230', 'R', '1231-1232', 'P', '1233-1234', 'K', '1235-1236', 'R', '1237-1238', 'P', '1239-1240', 'K', '1241-1242', 'R', '1243-1244', 'P', '1245-1246', 'K', '1247-1248', 'R', '1249-1250', 'P', '1251-1252', 'K', '1253-1254', 'R', '1255-1256', 'P', '1257-1258', 'K', '1259-1260', 'R', '1261-1262', 'P', '1263-1264', 'K', '1265-1266', 'R', '1267-1268', 'P', '1269-1270', 'K', '1271-1272', 'R', '1273-1274', 'P', '1275-1276', 'K', '1277-1278', 'R', '1279-1280', 'P', '1281-1282', 'K', '1283-1284', 'R', '1285-1286', 'P', '1287-1288', 'K', '1289-1290', 'R', '1291-1292', 'P', '1293-1294', 'K', '1295-1296', 'R', '1297-1298', 'P', '1299-1300', 'K', '1301-1302', 'R', '1303-1304', 'P', '1305-1306', 'K', '1307-1308', 'R', '1309-1310', 'P', '1311-1312', 'K', '1313-1314', 'R', '1315-1316', 'P', '1317-1318', 'K', '1319-1320', 'R', '1321-1322', 'P', '1323-1324', 'K', '1325-1326', 'R', '1327-1328', 'P', '1329-1330', 'K', '1331-1332', 'R', '1333-1334', 'P', '1335-1336', 'K', '1337-1338', 'R', '1339-1340', 'P', '1341-1342', 'K', '1343-1344', 'R', '1345-1346', 'P', '1347-1348', 'K', '1349-1350', 'R', '1351-1352', 'P', '1353-1354', 'K', '1355-1356', 'R', '1357-1358', 'P', '1359-1360', 'K', '1361-1362', 'R', '1363-1364', 'P', '1365-1366', 'K', '1367-1368', 'R', '1369-1370', 'P', '1371-1372', 'K', '1373-1374', 'R', '1375-1376', 'P', '1377-1378', 'K', '1379-1380', 'R', '1381-1382', 'P', '1383-1384', 'K', '1385-1386', 'R', '1387-1388', 'P', '1389-1390', 'K', '1391-1392', 'R', '1393-1394', 'P', '1395-1396', 'K', '1397-1398', 'R', '1399-1400', 'P', '1401-1402', 'K', '1403-1404', 'R', '1405-1406', 'P', '1407-1408', 'K', '1409-1410', 'R', '1411-1412', 'P', '1413-1414', 'K', '1415-1416', 'R', '1417-1418', 'P', '1419-1420', 'K', '1421-1422', 'R', '1423-1424', 'P', '1425-1426', 'K', '1427-1428', 'R', '1429-1430', 'P', '1431-1432', 'K', '1433-1434', 'R', '1435-1436', 'P', '1437-1438', 'K', '1439-1440', 'R', '1441-1442', 'P', '1443-1444', 'K', '1445-1446', 'R', '1447-1448', 'P', '1449-1450', 'K', '1451-1452', 'R', '1453-1454', 'P', '1455-1456', 'K', '1457-1458', 'R', '1459-1460', 'P', '1461-1462', 'K', '1463-1464', 'R', '1465-1466', 'P', '1467-1468', 'K', '1469-1470', 'R', '1471-1472', 'P', '1473-1474', 'K', '1475-1476', 'R', '1477-1478', 'P', '1479-1480', 'K', '1481-1482', 'R', '1483-1484', 'P', '1485-1486', 'K', '1487-1488', 'R', '1489-1490', 'P', '1491-1492', 'K', '1493-1494', 'R', '1495-1496', 'P', '1497-1498', 'K', '1499-1500', 'R', '1501-1502', 'P', '1503-1504', 'K', '1505-1506', 'R', '1507-1508', 'P', '1509-1510', 'K', '1511-1512', 'R', '1513-1514', 'P', '1515-1516', 'K', '1517-1518', 'R', '1519-1520', 'P', '1521-1522', 'K', '1523-1524', 'R', '1525-1526', 'P', '1527-1528', 'K', '1529-1530', 'R', '1531-1532', 'P', '1533-1534', 'K', '1535-1536', 'R', '1537-1538', 'P', '1539-1540', 'K', '1541-1542', 'R', '1543-1544', 'P', '1545-1546', 'K', '1547-1548', 'R', '1549-1550', 'P', '1551-1552', 'K', '1553-1554', 'R', '1555-1556', 'P', '1557-1558', 'K', '1559-1560', 'R', '1561-1562', 'P', '1563-1564', 'K', '1565-1566', 'R', '1567-1568', 'P', '1569-1570', 'K', '1571-1572', 'R', '1573-1574', 'P', '1575-1576', 'K', '1577-1578', 'R', '1579-1580', 'P', '1581-1582', 'K', '1583-1584', 'R', '1585-1586', 'P', '1587-1588', 'K', '1589-1590', 'R', '1591-1592', 'P', '1593-1594', 'K', '1595-1596', 'R', '1597-1598', 'P', '1599-1600', 'K', '1601-1602', 'R', '1603-1604', 'P', '1605-1606', 'K', '1607-1608', 'R', '1609-1610', 'P', '1611-1612', 'K', '1613-1614', 'R', '1615-1616', 'P', '1617-1618', 'K', '1619-1620', 'R', '1621-1622', 'P', '1623-1624', 'K', '1625-1626', 'R', '1627-1628', 'P', '1629-1630', 'K', '1631-1632', 'R', '1633-1634', 'P', '1635-1636', 'K', '1637-1638', 'R', '1639-1640', 'P', '1641-1642', 'K', '1643-1644', 'R', '1645-1646', 'P', '1647-1648', 'K', '1649-1650', 'R', '1651-1652', 'P', '1653-1654', 'K', '1655-1656', 'R', '1657-1658', 'P', '1659-1660', 'K', '1661-1662', 'R', '1663-1664', 'P', '1665-1666', 'K', '1667-1668', 'R', '1669-1670', 'P', '1671-1672', 'K', '1673-1674', 'R', '1675-1676', 'P', '1677-1678', 'K', '1679-1680', 'R', '1681-1682', 'P', '1683-1684', 'K', '1685-1686', 'R', '1687-1688', 'P', '1689-1690', 'K', '1691-1692', 'R', '1693-1694', 'P', '1695-1696', 'K', '1697-1698', 'R', '1699-1700', 'P', '1701-1702', 'K', '1703-1704', 'R', '1705-1706', 'P', '1707-1708', 'K', '1709-1710', 'R', '1711-1712', 'P', '1713-1714', 'K', '1715-1716', 'R', '1717-1718', 'P', '1719-1720', 'K', '1721-1722', 'R', '1723-1724', 'P', '1725-1726', 'K', '1727-1728', 'R', '1729-1730', 'P', '1731-1732', 'K', '1733-1734', 'R', '1735-1736', 'P', '1737-1738', 'K', '1739-1740', 'R', '1741-1742', 'P', '1743-1744', 'K', '1745-1746', 'R', '1747-1748', 'P', '1749-1750', 'K', '1751-1752', 'R', '1753-1754', 'P', '1755-1756', 'K', '1757-1758', 'R', '1759-1760', 'P', '1761-1762', 'K', '1763-1764', 'R', '1765-1766', 'P', '1767-1768', 'K', '1769-1770', 'R', '1771-1772', 'P', '1773-1774', 'K', '1775-1776', 'R', '1777-1778', 'P', '1779-1780', 'K', '1781-1782', 'R', '1783-1784', 'P', '1785-1786', 'K', '1787-1788', 'R', '1789-1790', 'P', '1791-1792', 'K', '1793-1794', 'R', '1795-1796', 'P', '1797-1798', 'K', '1799-1800', 'R', '1801-1802', 'P', '1803-1804', 'K', '1805-1806', 'R', '1807-1808', 'P', '1809-1810', 'K', '1811-1812', 'R', '1813-1814', 'P', '1815-1816', 'K', '1817-1818', 'R', '1819-1820', 'P', '1821-1822', 'K', '1823-1824', 'R', '1825-1826', 'P', '1827-1828', 'K', '1829-1830', 'R', '1831-1832', 'P', '1833-1834', 'K', '1835-1836', 'R', '1837-1838', 'P', '1839-1840', 'K', '1841-1842', 'R', '1843-1844', 'P', '1845-1846', 'K', '1847-1848', 'R', '1849-1850', 'P', '1851-1852', 'K', '1853-1854', 'R', '1855-1856', 'P', '1857-1858', 'K', '1859-1860', 'R', '1861-1862', 'P', '1863-1864', 'K', '1865-1866', 'R', '1867-1868', 'P', '1869-1870', 'K', '1871-1872', 'R', '1873-1874', 'P', '1875-1876', 'K', '1877-1878', 'R', '1879-1880', 'P', '1881-1882', 'K', '1883-1884', 'R', '1885-1886', 'P', '1887-1888', 'K', '1889-1890', 'R', '1891-1892', 'P', '1893-1894', 'K', '1895-1896', 'R', '1897-1898', 'P', '1899-1900', 'K', '1901-1902', 'R', '1903-1904', 'P', '1905-1906', 'K', '1907-1908', 'R', '1909-1910', 'P', '1911-1912', 'K', '1913-1914', 'R', '1915-1916', 'P', '1917-1918', 'K', '1919-1920', 'R', '1921-1922', 'P', '1923-1924', 'K', '1925-1926', 'R', '1927-1928', 'P', '1929-1930', 'K', '1931-1932', 'R', '1933-1934', 'P', '1935-1936', 'K', '1937-1938', 'R', '1939-1940', 'P', '1941-1942', 'K', '1943-1944', 'R', '1945-1946', 'P', '1947-1948', 'K', '1949-1950', 'R', '1951-1952', 'P', '1953-1954', 'K', '1955-1956', 'R', '1957-1958', 'P', '1959-1960', 'K', '1961-1962', 'R', '1963-1964', 'P', '1965-1966', 'K', '1967-1968', 'R', '1969-1970', 'P', '1971-1972', 'K', '1973-1974', 'R', '1975-1976', 'P', '1977-1978', 'K', '1979-1980', 'R', '1981-1982', 'P', '1983-1984', 'K', '1985-1986', 'R', '1987-1988', 'P', '1989-1990', 'K', '1991-1992', 'R', '1993-1994', 'P', '1995-1996', 'K', '1997-1998', 'R', '1999-2000', 'P', '2001-2002', 'K', '2003-2004', 'R', '2005-2006', 'P', '2007-2008', 'K', '2009-2010', 'R', '2011-2012', 'P', '2013-2014', 'K', '2015-2016', 'R', '2017-2018', 'P', '2019-2020', 'K', '2021-2022', 'R', '2023-2024', 'P', '2025-2026', 'K', '2027-2028', 'R', '2029-2030', 'P', '2031-2032', 'K', '2033-2034', 'R', '2035-2036', 'P', '2037-2038', 'K', '2039-2040', 'R', '2041-2042', 'P', '2043-2044', 'K', '2045-2046', 'R', '2047-2048', 'P', '2049-2050', 'K', '2051-2052', 'R', '2053-2054', 'P', '2055-2056', 'K', '2057-2058', 'R', '2059-2060', 'P', '2061-2062', 'K', '2063-2064', 'R', '2065-2066', 'P', '2067-2068', 'K', '2069-2070', 'R', '2071-2072', 'P', '2073-2074', 'K', '2075-2076', 'R', '2077-2078', 'P', '2079-2080', 'K', '2081-2082', 'R', '2083-2084', 'P', '2085-2086', 'K', '2087-2088', 'R', '2089-2090', 'P', '2091-2092', 'K', '2093-2094', 'R', '2095-2096', 'P', '2097-2098', 'K', '2099-2100', 'R', '2101-2102', 'P', '2103-2104', 'K', '2105-2106', 'R', '2107-2108', 'P', '2109-2110', 'K', '2111-2112', 'R', '2113-2114', 'P', '2115-2116', 'K', '2117-2118', 'R', '2119-2120', 'P', '2121-2122', 'K', '2123-2124', 'R', '2125-2126', 'P', '2127-2128', 'K', '2129-2130', 'R', '2131-2132', 'P', '2133-2134', 'K', '2135-2136', 'R', '2137-2138', 'P', '2139-2140', 'K', '2141-2142', 'R', '2143-2144', 'P', '2145-2146', 'K', '2147-2148', 'R', '2149-2150', 'P', '2151-2152', 'K', '2153-2154', 'R', '2155-2156', 'P', '2157-2158', 'K', '2159-2160', 'R', '2161-2162', 'P', '2163-2164', 'K', '2165-2166', 'R', '

Ri-Lee, S.; Greenspan, D. S.
Biochem. J. 310, 15-22, 1995
A;Title: Transcriptional promoter of the human alpha-1(V) collagen gene (COL5A1).
A;Reference number: S58665; MUID:9537437
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-36 <LEE>
A;Cross-references: GB:L38808; NID:91020325; PIDN:AAA79853.1; PID:91020326
C;Comment: Prolines and lysines at the third position of the tripeptide repeating unit are 5-hydroxylated and subsequently O-glycosylated.
C;Comment: A long form of the mature protein containing the heterotrimeric part of the amino-terminal propeptide the heterotrimmers are probably processed to the long form.
C;Genetics: GDB:COL5A1
A;Map position: GDB:131457; OMIM:120215
A;Cross-references: GDB:131457; OMIM:120215
C;Complex: type V collagen may be a homotrimer of alpha 1(V) chains, a heterotrimer of alpha 2(V) chain and one alpha 3(V) chain, initially linked by disulfide bonds among the chains, is formed with desmosine cross-links made from lysine and allysine residues
C;Description: structural component of extracellular fibrous polymer associated with cell
C;Superfamily: collagen I fibrils
C;Keywords: coiled coil; extracellular matrix; glycoprotein; hydroxylysine; hydroxyproline
F;1-37/Domain: signal sequence #status predicted <SIG>
F;36-261/Domain: PAR1-like #status predicted <PAR>
F;38-51/Domain: aminoterminal propeptide #status predicted <PRO>
F;542-1605/Product: collagen alpha 1(V) chain, short form #status predicted <MAT>
F;559-1572/Region: aminoterminal nonhelical telopeptide
F;456-647/Region: cell attachment (R-G-D) motif
F;663-665/Region: heparin binding
F;1573-1601/Region: carboxyl-terminal nonhelical telopeptide
F;891-929/Region: carboxyl-terminal nonhelical telopeptide
F;1600-1834/Region: carboxyl-terminal nonhelical telopeptide
F;1615-1837/Region: fibrillar collagen carboxyl-terminal homology <FCC>
F;1620-1842/Region: fibrillar collagen carboxyl-terminal homology <FCC>
F;162-244/Region: pyrrolidine carboxylic acid (Gln) (in mature form) #status predicted
F;159-176,385-1672,171/Region: disulfide bonds: #status predicted
F;234-235,240-262,263-273,274-275,277-279,280-338,340-346,347,352,357,416,417,420,421/Region: modified site: allysine (Lys) #status predicted
F;541-542/Cleavage site: Ala-Gln (procollagen N-endopeptidase) #status predicted
F;542/Modified site: pyrrolidine carboxylic acid (Gln) (in mature form) #status predicted
F;570,576-621,639-648,654-657,675,678,690,693-696,705,717,720,726,732,741,750,753-756,78/Region: 4-hydroxyproline (Pro) #status experimental
F;627,642,687,708,744-774,795,804,807,810,819,825,846,864,882,897/Region: modified site: 5-hydroxylysine (Lys) (covalent) #status experimental
F;708,744/Region: binding site: carbohydrate (Lys) (covalent) #status experimental
F;1482/Modified site: 5-hydroxylysine (Lys) #status predicted
F;1639,1645,1662,1671/Region: disulfide bonds: interchain #status predicted
F;1680-1835,1746-1789/Region: disulfide bonds: #status predicted

Query Match 19.3%; Score 479.5; DB 2; Length 1843;
Best Local Similarity 44.1%;保守性 9; Mismatches 9; Pred. No. 7e-22; Indels 57; Gaps 4;
Matches 105; Conservative 9; Mismatches 67; Indels 57; Gaps 4;

Qy 41 P^TE^NG^IL^PG^RD^RG^RG^PR^EK^GD^GL^PG^MG^LS^GL^QG^TP^GP^GK^GE^NS^AC^EP^GP^KE 100
Db 1323 P^TC^PG^RP^GD^DC^PK^GS^PG^PV^FG^PG^PP^ER^PG^SQ^DG^PD^DG^ER^GP^TS^GP^GT^E 1382
Qy 101 R^GL^SG^PG^LC^IP^GA^CK⁻⁻⁻EG^SA^RQ^NG^PQ⁻⁻⁻E^GP^SA^RQ^NG^PQ⁻⁻⁻E^GK^GP⁻⁻⁻ 136
Db 1388 P^GS^PG^PG^KK^GP^GP^GA^CG^PE^RQ^GR^EG^KA^GE^AL^EG^LG^PC^KT⁻⁻⁻G^PQ^GA^PG^RF^GP^DG^LR^I 1447
Qy 137 -----K^GA^GP^KE^VG^AP^GM^OO^ST^GA^KG^ST^PK^GE 166
Db 1448 P^GP^VE^QG^IP^GS^PG^PD^GR^PP^GM^GP^GL^IK^GD^SP^GK^EK^GH^GL^IG^LG^IG^PG^EK^GD 1507
Qy 167 R^GA^PV^OG^A---P^NA^GA^GP^AG^PG^AG^PG^AS^SR^GP^GL^KD^DR^GV^GP^DR^GK^IG^ES^LP 221
Db 1508 R^GL^PG^PO^SS^GP^KE^QG^IT^GP^SL^GP^FG^PG^PG^KA^GS^SG^PT^GP^KE^AG^HP 1565

Search completed: July 3, 2001, 16:16:42
Job time: 159 sec

Db 1503 R^GL^PG^PQ^GS^SG^PK^GE^QG^IT^GP^SG^PI^GP^GP^GG^LP^GP^GK^GA^KG^SS^GT^GP^KE^AG^HP 1560
RESULT 15
S18803 collagen alpha 1(V) chain - hamster
C;Species: Cricetinae gen. sp. (hamster)
C;Date: 19-Mar-1997 #sequence_revision 24-Jul-1997 #text_change 16-Dec-1998
C;Accession: S18803
R;Greenbaum, D.S.; Cheng, W.; Hoffman, G.G.
J. Biol. Chem. 266, 2477-2477, 1991
A;Title: The pro-alpha1(V) collagen chain. Complete primary structure, distribution and properties
A;Reference number: S18802; MUID:92105142
A;Accession: S18803
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1843 <GRE>
C;Superfamily: collagen alpha 1(V) chain; fibrillar collagen carboxyl-terminal homology <FCC>
F;1620-1842/Region: fibrillar collagen carboxyl-terminal homology <FCC>

Qy 41 P^TE^NG^IL^PG^RD^RG^RG^PR^EK^GD^GL^PG^MG^LS^GL^QG^TP^GP^GK^GE^NS^AC^EP^GP^KE 100
Db 1328 P^TP^GK^GP^GD^DC^PK^GS^PG^PV^FG^PG^PP^ER^PG^SQ^DG^PD^DG^ER^GP^TS^GP^GT^E 1387
Qy 101 R^GL^SG^PG^LC^IP^GA^CK⁻⁻⁻EG^SA^RQ^NG^PQ⁻⁻⁻E^GP^SA^RQ^NG^PQ⁻⁻⁻E^GK^GP⁻⁻⁻ 136
Db 1388 P^GS^PG^PG^KK^GP^GP^GA^CG^PE^RQ^GR^EG^KA^GE^AL^EG^LG^PC^KT⁻⁻⁻G^PQ^GA^PG^RF^GP^DG^LR^I 1447
Qy 137 -----K^GA^GP^KE^VG^AP^GM^OO^ST^GA^KG^ST^PK^GE 166
Db 1448 P^GP^VE^QG^IP^GS^PG^PD^GR^PP^GM^GP^GL^IK^GD^SP^GK^EK^GH^GL^IG^LG^IG^PG^EK^GD 1507
Qy 167 R^GA^PV^OG^A---P^NA^GA^GP^AG^PG^AG^PG^AS^SR^GP^GL^KD^DR^GV^GP^DR^GK^IG^ES^LP 221
Db 1508 R^GL^PG^PO^SS^GP^KE^QG^IT^GP^SL^GP^FG^PG^PG^KA^GS^SG^PT^GP^KE^AG^HP 1565

GenCore version 4.5
 copyright (c) 1993 - 2000 Compugen Ltd.
 OM protein - protein search, using SW model
 Run on: July 3, 2001, 16:15:03 : Search time 12:03 Seconds
 (without alignments)
 1341.175 Million cell updates/sec

Title: US-09-454-223-2

Perfect score: 2487

Sequence: 1 MLPFLSMLVIVLVLVQNLGA.....TDPSSQVSHGTGFTSFGLLKL 471

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SwissProt_39: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	1384.5	55.7	374	PSPD_MOUSE
2	1286	51.7	374	PSPD_RAT
3	1110	44.6	261	1 TNF5_HUMAN
4	1053	42.3	375	1 PSPD_BOVIN
5	974	39.2	261	1 PSPD_BOVIN
6	967.5	38.9	369	1 TNF5_BOVIN
7	939.5	37.8	260	1 TNF5_BOVINA
8	907.5	36.5	260	1 TNF5_CANFA
9	878	35.3	371	1 CONG_BOVIN
10	827	33.3	260	1 TNF5_MOUSE
11	565.5	22.7	301	1 CLA3_BOVIN
12	484.5	19.5	1049	1 CA13_BOVIN
13	482.5	19.4	1460	1 CA11_CANFA
14	481.5	19.4	1464	1 CA11_HUMAN
15	481	19.3	1453	1 CA11_CHICK
16	480.5	19.3	1804	1 CA1B_MOUSE
17	480	19.3	1806	1 CA1B_HUMAN
18	478	19.2	1650	1 CA2B_MOUSE
19	477.5	19.2	1464	1 CA13_MOUSE
20	476.5	19.2	1453	1 CA11_MOUSE
21	476.5	19.2	1838	1 CA15_HUMAN
22	475	19.1	671	1 CA11_RAT
23	474.5	19.1	1466	1 CA13_HUMAN
24	474	19.1	1736	1 CA2B_HUMAN
25	473.5	19.0	1418	1 CA12_HUMAN
26	472.5	19.0	779	1 CA11_BOVIN
27	472.5	19.0	1763	1 CA24_ASCSU
28	470.5	18.9	911	1 CA1B_BOVIN
29	470	18.9	1022	1 CA26_CHICK
30	469.5	18.9	1459	1 CA12_MOUSE
31	469.5	18.9	2944	1 CA17_HUMAN
32	468.5	18.8	482	1 CA1B_RAT
33	462.5	18.6	2039	1 CA12_BOVIN

RESULT	1	ALIGNMENTS
PSPD_MOUSE		
ID: PSPD_MOUSE	STANDARD:	PRT: 374 AA.
AC: P50404;		
DT: 01-OCT-1996 (Rel. 34, Created)		
DT: 01-OCT-1996 (Rel. 34, Last sequence update)		
DT: 15-JUL-1999 (Rel. 38, Last annotation update)		
DE: PULMONARY SURFACTANT-ASSOCIATED PROTEIN D PRECURSOR (SP-D) (PSP-D)		
GN: SFTPD OR SFTPA		
OS: Mus musculus (Mouse)		
OC: Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;		
OC: Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus		
OX: NCBI_TaxID-10090;		
RN: [1] J. Immunol. 155:5671-5677(1995).		
RP: SEQUENCE FROM N.A.		
RC: STRAIN-B6CBA; TISSUE-Lung;		
RX: MEDLINE-96094460; PUBMED-7499852;		
RA: Motwani M., White R.A., Guo N., Dowler L.L., Tauber A.I., Sastry K.N.,		
RT: "Mouse surfactant protein-D, cDNA cloning, characterization, and gene localization to chromosome 14."		
RL: J. Immunol. 155:5671-5677(1995).		
RP: SEQUENCE FROM N.A.		
CC: MICROORGANISMS, LINDS STRONGLY MALTOSE RESIDUES AND TO A LESSER EXTENT OTHER ALPHA-GLUCOSYL MOIETIES. IT COULD PARTICIPATE IN THE EXTRACELLULAR REORGANIZATION OR TURNOVER OF PULMONARY SURFACTANT.		
CC: SUBUNIT: OLIGOMERIC COMPLEX OF 4 SET OF HOMOTRIMERS.		
CC: SUBCELLULAR LOCATION: EXTRACELLULAR.		
CC: MISCELLANOUS: PULMONARY SURFACTANT CONSISTS OF 90% LIPID AND 10% PROTEIN. THERE ARE 4 SURFACTANT ASSOCIATED PROTEIN: 2 COLLAGENOUS, HYDROPHOBIC-BANDING GLYCOPROTEINS (SP-A AND SP-D) AND 2 SMALL		
CC: HYDROPHOBIC PROTEINS (SP-B AND SP-C).		
CC: -1- SIMILARITY: CONTAINS A COLLAGENOUS DOMAIN.		
CC: -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.		
CC: This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).		
CC: EMBL: LA0156; AAA9021.1; -.		
DR: HSSP: P19999; 1MSB.		
DR: MGII: 109515; Sftp.		
DR: InterPro: IPR00087; -.		
DR: InterPro: IPR01304; -.		
DR: Pfam: PF01391; Collagen_3.		
DR: Pfam: PF0059; lectinc_1.		
DR: PROSITE: PS00615; C_TYPELECTIN_1; -.		
DR: PROSITE: PS50041; C_TYPELECTIN_2; 1.		
KW: Glycoprotein; Calcium; Surface film; Gaseous exchange; Hydroxylation; Signal; Lectin; Collagen; Repeat; Coiled coil.		
FT: SIGNAL	19	BY SIMILARITY.
FT: CHAIN	20	PULMONARY SURFACTANT-ASSOCIATED PROTEIN
FT: 374	D.	

DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE CD40 LIGAND (CD40-L) (TNF-RELATED ACTIVATION PROTEIN) (TRAP) (T CELL ANTIGEN GP39) (CD154 ANTIGEN)
 DE TNFSE5 OR CD40LG OR CD40L OR TRAP
 GN Homo sapiens (Human)
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
 OX NCBI_TAXID=9606;
 RN [1] SEQUENCE FROM N.A.
 RX MEDLINE=93076554; PubMed=1280226;
 RA Graf D., Korttneuer U., Mages H.W., Senger G., Kroczeck R.A.;
 RT "Cloning of TRAP, a ligand for CD40 on human T cells.";
 RL Eur. J. Immunol. 22:3191-3194(1992).
 RN [2] SEQUENCE FROM N.A.
 RX MEDLINE=93049181; PubMed=1385114;
 RA Hollenbaugh D., Grossmaire L.S., Kullas C.D., Chalupny J.N.,
 RA Brasch-Andersen S., Nobile R.J., Stamenkovic I., Ledbetter J.A.,
 RA Aruffo A.;
 RT "The human T cell antigen gp39, a member of the TNF gene family, is a ligand for the CD40 receptor: expression of a soluble form of gp39 with B cell co-stimulatory activity.";
 RT EMBO J. 11:4313-4321(1992).
 RN [3] SEQUENCE FROM N.A. AND VARIANTS HIGM1 128-ARG-GLY-129 AND PRO-235.
 RX MEDLINE=93145330; PubMed=7618782;
 RA Aruffo A., Farrington M., Hollenbaugh D., Li X., Milatovich A.,
 RA Nonoyama S., Bajorth J., Grossmaire L.S., Stenkamp R., Neubauer M.,
 RA Roberts R.L., Nobile R.J., Ledbetter J.A., Francke U., Ochs H.D.,
 RT "The CD40 ligand, gp39, is defective in activated T cells from patients with X-linked hyper-IgM syndrome.";
 RL Cell 72:291-300(1993).
 RN [4] SEQUENCE FROM N.A.
 RX MEDLINE=93094157; PubMed=12811209;
 RA Spriggs M.K., Armitage R.J., Strockbine L., Clifford K.N.,
 RA Macduff B.M., Sato T.A., Maliszewski C.R., Fanslow W.C.;
 RT "Recombinant human CD40 ligand stimulates B cell proliferation and immunoglobulin E secretion.";
 RT J. Exp. Med. 176:1543-1550(1992).
 RN [5] SEQUENCE FROM N.A.
 RX MEDLINE=93138085; PubMed=7670552;
 RA Gauchat J.F.M., Aubry J., Mazzel G.J., Life P., Jomotte T., Elson G.,
 RA "Human CD40-ligand: molecular cloning, cellular distribution and regulation of expression by factors controlling IgE production.";
 RL FEBS Lett. 315:25-28(1993).
 RN [6] SEQUENCE FROM N.A.
 RX Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
 RN [7] X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 116-261.
 RX MEDLINE=661131874; PubMed=8588998;
 RA Karplusas M., Hsu Y.-M., Wang J.-H., Thompson J., Lederman S.,
 RA Chees L., Thomas D.;
 RA "2-A crystal structure of an extracellular fragment of human CD40 ligand.";
 RL Structure 3:1031-1039(1995).
 RN [8] 3D-STRUCTURE MODELING OF COMPLEX WITH CD40.
 RX MEDLINE=98266353; PubMed=9605317;
 RA Singh J., Garber E., van Viljtem H., Karpusas M., Hsu Y.-M.,
 RA Zheng Z., Naimi H., Thomas D.,
 RT "The role of polar interactions in the molecular recognition of CD40L with its receptor CD40.";
 RL Protein Sci. 7:1124-1135(1998).
 RN [9]

RP VARIANTS HIGM1 ARG-36 AND GLY-140.
 RX MEDLINE=93156339; PubMed=7679206;
 RA Korttneuer U., Graf D., Mages H.W., Briere F., Padayachee M.,
 RA Malcolm S., Uazio A.G., Notarangelo L.D., Levinsky R.J.,
 RA Kroczeck R.A.;
 RT "Defective expression of T-cell CD40 ligand causes X-linked immunodeficiency with hyper-IgM.";
 RL Nature 361:539-541(1993).
 RN [10] VARIANT HIGM1 GLU-123.
 RX MEDLINE=93156340; PubMed=8094231;
 RA D'Antonio J.P., Bonnaffoy J.Y., Gauchat J.F.M., Fischer A.,
 RA de Saint Basile G.;
 RT "CD40 ligand mutations in x-linked immunodeficiency with hyper-IgM.";
 RL Nature 361:541-543(1993).
 RN [11] VARIANTS HIGM1 PRO-155; ASP-211 AND VAL-227.
 RX MEDLINE=93174270; PubMed=7679801;
 RA Allen R.C., Armitage R.J., Conley M.E., Rosenblatt H., Jenkins N.A.,
 RA Copeland N.G., Bedell M.A., Bedell S., Distefano C.M.,
 RA Simmeaux D.K., Fanslow W.C., Belmont J.W., Spriggs M.K.;
 RT "CD40 ligand gene defects responsible for X-linked hyper-IgM syndrome.";
 RL Science 259:980-993(1993).
 RN [12] VARIANTS HIGM1 ALA-126; ARG-140 AND GLU-144.
 RX MEDLINE=95233138; PubMed=7717401;
 RA Macchi P., Villa A., Stolina D., Sacco M.G., Morall F., Brugnoni D.,
 RA Gilliani S., Mantuano E., Fasth A., Andersson B., Zegers B.J.M.,
 RA Cavigni G., Mantuano E., Fasth A., Andersson B., Zegers B.J.M.,
 RA Piebani A., Vezzoni P., Notarangelo L.D., Poret Y., Airo P.,
 RT "Characterization of nine novel mutations in the CD40 ligand gene in patients with X-linked hyper IgM syndrome of various ancestry.";
 RL Am. J. Hum. Genet. 56:888-906(1995).
 RN [13] VARIANTS HIGM1 PRO-155 AND VAL-227, AND VARIANT ARG-219.
 RX MEDLINE=96133533; PubMed=8550833;
 RA Lin Q., Rohrer J., Allen R.C., Laiche M., Greene J.M., Shigeoka A.O.,
 RA Gatti R.A., Deauf D.C., Belmont J.W., Conley M.E.;
 RT "A single strand conformation polymorphism study of CD40 ligand: Efficient mutation analysis and carrier detection for X-linked hyper IgM syndrome.";
 RL J. Clin. Invest. 97:196-201(1996).
 RN [14] VARIANTS HIGM1 ARG-36; CYS-140; SR-231; MET-254 AND GLY-227 DEL.
 RX MEDLINE=97295077; PubMed=9150729;
 RA Nonoyama S., Shimadzu M., Toru H., Seyama K., Nuno H., Neubauer M.,
 RA Yata J.-I., Ochi H.D.;
 RT "Mutations of the CD40 ligand gene in 13 Japanese patients with X-linked hyper-IgM syndrome.";
 RL X-Linked Genet. 93:624-627(1997).
 CC -1- FUNCTION: MEDIATES B-CELL PROLIFERATION IN THE ABSENCE OF CO-STIMULUS AS WELL AS IgE PRODUCTION IN THE PRESENCE OF IL-4.
 CC -1- DISEASE: DEFECTS IN CD40LG ARE THE CAUSE OF AN X-LINKED INVOLVED IN IMMUNOGLOBULIN CLASS SWITCHING.
 CC -1- SUBUNIT: HOMOTRIMER.
 CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN EXTRACELLULAR SOLUBLE FORM.
 CC -1- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED ON ACTIVATED CD4+ T-LYMPHOCYTES.
 CC -1- DISEASE: DEFECTS IN CD40LG ARE THE CAUSE OF AN X-LINKED IMMUNODEFICIENCY WITH HYPER-IgM (HIGM1), AN IMMUNOGLOBULIN ISOTYPE SWITCH DEFECT CHARACTERIZED BY Elevated CONCENTRATIONS OF SERUM IgM AND DECREASED AMOUNTS OF ALL OTHER ISOTYPES. AFFECTED MALES PRESENT AT AN EARLY AGE (USUALLY WITHIN THE FIRST YEAR OF LIFE) RECURRENT BACTERIAL AND OPPORTUNISTIC INFECTIONS, INCLUDING PNEUMOCYSTIS CARINII PNEUMONIA AND INTRACRANIAL DIARRHEA DUE TO CRYPTOSPORIDIUM INFECTION. DESPITE SUBSTITUTIVE TREATMENT WITH INTRAVENOUS IMMUNOGLOBULIN, THE OVERALL PROGNOSIS IS RATHER POOR, WITH A DEATH RATE OF ABOUT 10% BEFORE ADOLESCENCE.
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
 CC -1- DATABASE: NAME=CD40Database;
 CC NOTE=European CD40L defect database (mutation db);
 CC WWW="http://www.exasy.ch/cd40lbase/";

DR PDB: 1B0B; 29-NOV-99.
 DR MIM: 176635; -.
 DR InterPro: IPR000087; -.
 DR Pfam: PR01391; Collagen; 3.
 DR PROSITE: PS00615; C-TYPE_LECTIN_1.
 DR PROTEIN: PS50041; C-TYPE_LECTIN_2; 1.
 KW Glycoprotein; Calcium; Surface film; Gaseous exchange; Hydroxylation; Signal; Lectin; Collagen; Repeat; Coiled coil; 3D-structure.
 PT SIGNAL 1 20 BY SIMILARITY.
 FT CHAIN 21 375 PULMONARY-SURFACTANT-ASSOCIATED PROTEIN D.
 FT DISULFID 281 373
 FT DISULFID 351
 FT MOD_RES 721 721 BLOCKED.
 FT CARBOHYD 90 90 N-LINKED (GLCNAC. - -) (POTENTIAL).
 FT MOD_RES 78 78 HYDROXYLATION (BY SIMILARITY).
 FT MOD_RES 87 87 HYDROXYLATION (BY SIMILARITY).
 FT MOD_RES 96 96 HYDROXYLATION (BY SIMILARITY).
 FT MOD_RES 99 99 HYDROXYLATION (BY SIMILARITY).
 FT MOD_RES 171 171 HYDROXYLATION (BY SIMILARITY).
 FT MOD_RES 177 177 HYDROXYLATION (BY SIMILARITY).
 FT CONFLICT 31 31 M -> T (IN REF. 2).
 FT CONFLICT 59 59 P -> F (IN REF. 3).
 FT CONFLICT 122 122 A -> P (IN REF. 2).
 FT CONFLICT 180 180 T -> A (IN REF. 2).
 FT CONFLICT 206 206 D -> P (IN REF. 3).
 FT CONFLICT 374 374 E -> EH (IN REF. 3).
 SQ SEQUENCE 375 AA: 37702 MW: 2986B2699FC01A6A CRC64;

Query Match 42.3% Score 1053; DB 1; Length 375; Best Local Similarity 72.6%; Pred. No. 6.7e-52; Mismatches 191; Conservative 28; Mismatches 44; Indels 0; Gaps 0; CC
 FT DOMAIN 46 224
 FT DISULFID 46 224
 FT TRANSMEM 23 46
 FT DOMAIN 1 22
 FT CARBOHYD 178 218
 FT SEQUENCE 261 AA: 29242 MW: 8491FEB30A787FD CRC64;

Query Match 39.2% Score 974; DB 1; Length 261; Best Local Similarity 88.4%; Pred. No. 1.1e-47; Mismatches 190; Conservative 5; Mismatches 20; Indels 0; Gaps 0; CC
 FT DOMAIN 47 261 EXTRACELLULAR (POTENTIAL).
 FT DISULFID 178 218 N-LINKED (GLCNAC. - -) (POTENTIAL).
 FT CARBOHYD 240 240
 FT SEQUENCE 261 AA: 29242 MW: 8491FEB30A787FD CRC64;

Query Match 37.7% Score 914; DB 1; Length 261; Best Local Similarity 88.4%; Pred. No. 1.1e-47; Mismatches 190; Conservative 5; Mismatches 20; Indels 0; Gaps 0; CC
 FT DOMAIN 47 261 EXTRACELLULAR (POTENTIAL).
 FT DISULFID 178 218 N-LINKED (GLCNAC. - -) (POTENTIAL).
 FT CARBOHYD 240 240
 FT SEQUENCE 261 AA: 29242 MW: 8491FEB30A787FD CRC64;

RESULT 5
 TNF5_BOVIN ID TNF5_BOVIN STANDARD: PRT; 261 AA.
 AC P51749; DT 01-OCT-1995 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DE 30-MAY-2000 (Rel. 39, Last annotation update)
 DE CD40 LIGAND (TNF-RELATED ACTIVATION PROTEIN) (TRAP) (T CELL ANTIGEN DE GP99).
 DE TNFSF5 OR CD40LG OR CD40L.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; OC

OC Bovidae; Bovinae; Bos.
 OX NCBI_TAXID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Blood;
 RX MEDLINE-9600582; PubMed-7500981;
 RA Merrens B.E.L.C.; Muriuki M.; Gaidulis L.;
 RT "Cloning of two members of the TNF-superfamily in cattle: CD40 ligand and tumor necrosis factor alpha.";
 RL Immunogenetics 42:430-431(1995).
 CC -1- FUNCTION: MEDIATES B-CELL PROLIFERATION IN THE ABSENCE OF CO-STIMULUS AS WELL AS IGE PRODUCTION IN THE PRESENCE OF IL-4.
 CC INVOLVED IN IMMUNOCOLOBLIN CLASS SWITCHING (BY SIMILARITY).
 CC -1- SUBUNIT: HOMOTRIMER.
 CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN EXTRACELLULAR SOLUBLE FORM (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.ebi.sib.ch/announce/> or send an email to license@ebi.sib.ch).
 CC DR EMBL; 248469; CAB88363.1; -.
 DR HSSP; P29965; ITALY.
 DR InterPro: IPR000478; -.
 DR Pfam: PF00220; TNF_1; -.
 DR PROSITE: PS50049; TNF_2; 1.
 DR PROTEIN: PS00251; TNF_3; 1.
 DR PROTEIN: PS50049; TNF_4; 1.
 DR PROTEIN: PS50049; TNF_5; 1.
 DR PROTEIN: PS50049; TNF_6; 1.
 DR PROTEIN: PS50049; TNF_7; 1.
 DR PROTEIN: PS50049; TNF_8; 1.
 DR PROTEIN: PS50049; TNF_9; 1.
 DR PROTEIN: PS50049; TNF_10; 1.
 DR PROTEIN: PS50049; TNF_11; 1.
 DR PROTEIN: PS50049; TNF_12; 1.
 DR PROTEIN: PS50049; TNF_13; 1.
 DR PROTEIN: PS50049; TNF_14; 1.
 DR PROTEIN: PS50049; TNF_15; 1.
 DR PROTEIN: PS50049; TNF_16; 1.
 DR PROTEIN: PS50049; TNF_17; 1.
 DR PROTEIN: PS50049; TNF_18; 1.
 DR PROTEIN: PS50049; TNF_19; 1.
 DR PROTEIN: PS50049; TNF_20; 1.
 DR PROTEIN: PS50049; TNF_21; 1.
 DR PROTEIN: PS50049; TNF_22; 1.
 DR PROTEIN: PS50049; TNF_23; 1.
 DR PROTEIN: PS50049; TNF_24; 1.
 DR PROTEIN: PS50049; TNF_25; 1.
 DR PROTEIN: PS50049; TNF_26; 1.
 DR PROTEIN: PS50049; TNF_27; 1.
 DR PROTEIN: PS50049; TNF_28; 1.
 DR PROTEIN: PS50049; TNF_29; 1.
 DR PROTEIN: PS50049; TNF_30; 1.
 DR PROTEIN: PS50049; TNF_31; 1.
 DR PROTEIN: PS50049; TNF_32; 1.
 DR PROTEIN: PS50049; TNF_33; 1.
 DR PROTEIN: PS50049; TNF_34; 1.
 DR PROTEIN: PS50049; TNF_35; 1.
 DR PROTEIN: PS50049; TNF_36; 1.
 DR PROTEIN: PS50049; TNF_37; 1.
 DR PROTEIN: PS50049; TNF_38; 1.
 DR PROTEIN: PS50049; TNF_39; 1.
 DR PROTEIN: PS50049; TNF_40; 1.
 DR PROTEIN: PS50049; TNF_41; 1.
 DR PROTEIN: PS50049; TNF_42; 1.
 DR PROTEIN: PS50049; TNF_43; 1.
 DR PROTEIN: PS50049; TNF_44; 1.
 DR PROTEIN: PS50049; TNF_45; 1.
 DR PROTEIN: PS50049; TNF_46; 1.
 DR PROTEIN: PS50049; TNF_47; 1.
 DR PROTEIN: PS50049; TNF_48; 1.
 DR PROTEIN: PS50049; TNF_49; 1.
 DR PROTEIN: PS50049; TNF_50; 1.
 DR PROTEIN: PS50049; TNF_51; 1.
 DR PROTEIN: PS50049; TNF_52; 1.
 DR PROTEIN: PS50049; TNF_53; 1.
 DR PROTEIN: PS50049; TNF_54; 1.
 DR PROTEIN: PS50049; TNF_55; 1.
 DR PROTEIN: PS50049; TNF_56; 1.
 DR PROTEIN: PS50049; TNF_57; 1.
 DR PROTEIN: PS50049; TNF_58; 1.
 DR PROTEIN: PS50049; TNF_59; 1.
 DR PROTEIN: PS50049; TNF_60; 1.
 DR PROTEIN: PS50049; TNF_61; 1.
 DR PROTEIN: PS50049; TNF_62; 1.
 DR PROTEIN: PS50049; TNF_63; 1.
 DR PROTEIN: PS50049; TNF_64; 1.
 DR PROTEIN: PS50049; TNF_65; 1.
 DR PROTEIN: PS50049; TNF_66; 1.
 DR PROTEIN: PS50049; TNF_67; 1.
 DR PROTEIN: PS50049; TNF_68; 1.
 DR PROTEIN: PS50049; TNF_69; 1.
 DR PROTEIN: PS50049; TNF_70; 1.
 DR PROTEIN: PS50049; TNF_71; 1.
 DR PROTEIN: PS50049; TNF_72; 1.
 DR PROTEIN: PS50049; TNF_73; 1.
 DR PROTEIN: PS50049; TNF_74; 1.
 DR PROTEIN: PS50049; TNF_75; 1.
 DR PROTEIN: PS50049; TNF_76; 1.
 DR PROTEIN: PS50049; TNF_77; 1.
 DR PROTEIN: PS50049; TNF_78; 1.
 DR PROTEIN: PS50049; TNF_79; 1.
 DR PROTEIN: PS50049; TNF_80; 1.
 DR PROTEIN: PS50049; TNF_81; 1.
 DR PROTEIN: PS50049; TNF_82; 1.
 DR PROTEIN: PS50049; TNF_83; 1.
 DR PROTEIN: PS50049; TNF_84; 1.
 DR PROTEIN: PS50049; TNF_85; 1.
 DR PROTEIN: PS50049; TNF_86; 1.
 DR PROTEIN: PS50049; TNF_87; 1.
 DR PROTEIN: PS50049; TNF_88; 1.
 DR PROTEIN: PS50049; TNF_89; 1.
 DR PROTEIN: PS50049; TNF_90; 1.
 DR PROTEIN: PS50049; TNF_91; 1.
 DR PROTEIN: PS50049; TNF_92; 1.
 DR PROTEIN: PS50049; TNF_93; 1.
 DR PROTEIN: PS50049; TNF_94; 1.
 DR PROTEIN: PS50049; TNF_95; 1.
 DR PROTEIN: PS50049; TNF_96; 1.
 DR PROTEIN: PS50049; TNF_97; 1.
 DR PROTEIN: PS50049; TNF_98; 1.
 DR PROTEIN: PS50049; TNF_99; 1.
 DR PROTEIN: PS50049; TNF_100; 1.
 DR PROTEIN: PS50049; TNF_101; 1.
 DR PROTEIN: PS50049; TNF_102; 1.
 DR PROTEIN: PS50049; TNF_103; 1.
 DR PROTEIN: PS50049; TNF_104; 1.
 DR PROTEIN: PS50049; TNF_105; 1.
 DR PROTEIN: PS50049; TNF_106; 1.
 DR PROTEIN: PS50049; TNF_107; 1.
 DR PROTEIN: PS50049; TNF_108; 1.
 DR PROTEIN: PS50049; TNF_109; 1.
 DR PROTEIN: PS50049; TNF_110; 1.
 DR PROTEIN: PS50049; TNF_111; 1.
 DR PROTEIN: PS50049; TNF_112; 1.
 DR PROTEIN: PS50049; TNF_113; 1.
 DR PROTEIN: PS50049; TNF_114; 1.
 DR PROTEIN: PS50049; TNF_115; 1.
 DR PROTEIN: PS50049; TNF_116; 1.
 DR PROTEIN: PS50049; TNF_117; 1.
 DR PROTEIN: PS50049; TNF_118; 1.
 DR PROTEIN: PS50049; TNF_119; 1.
 DR PROTEIN: PS50049; TNF_120; 1.
 DR PROTEIN: PS50049; TNF_121; 1.
 DR PROTEIN: PS50049; TNF_122; 1.
 DR PROTEIN: PS50049; TNF_123; 1.
 DR PROTEIN: PS50049; TNF_124; 1.
 DR PROTEIN: PS50049; TNF_125; 1.
 DR PROTEIN: PS50049; TNF_126; 1.
 DR PROTEIN: PS50049; TNF_127; 1.
 DR PROTEIN: PS50049; TNF_128; 1.
 DR PROTEIN: PS50049; TNF_129; 1.
 DR PROTEIN: PS50049; TNF_130; 1.
 DR PROTEIN: PS50049; TNF_131; 1.
 DR PROTEIN: PS50049; TNF_132; 1.
 DR PROTEIN: PS50049; TNF_133; 1.
 DR PROTEIN: PS50049; TNF_134; 1.
 DR PROTEIN: PS50049; TNF_135; 1.
 DR PROTEIN: PS50049; TNF_136; 1.
 DR PROTEIN: PS50049; TNF_137; 1.
 DR PROTEIN: PS50049; TNF_138; 1.
 DR PROTEIN: PS50049; TNF_139; 1.
 DR PROTEIN: PS50049; TNF_140; 1.
 DR PROTEIN: PS50049; TNF_141; 1.
 DR PROTEIN: PS50049; TNF_142; 1.
 DR PROTEIN: PS50049; TNF_143; 1.
 DR PROTEIN: PS50049; TNF_144; 1.
 DR PROTEIN: PS50049; TNF_145; 1.
 DR PROTEIN: PS50049; TNF_146; 1.
 DR PROTEIN: PS50049; TNF_147; 1.
 DR PROTEIN: PS50049; TNF_148; 1.
 DR PROTEIN: PS50049; TNF_149; 1.
 DR PROTEIN: PS50049; TNF_150; 1.
 DR PROTEIN: PS50049; TNF_151; 1.
 DR PROTEIN: PS50049; TNF_152; 1.
 DR PROTEIN: PS50049; TNF_153; 1.
 DR PROTEIN: PS50049; TNF_154; 1.
 DR PROTEIN: PS50049; TNF_155; 1.
 DR PROTEIN: PS50049; TNF_156; 1.
 DR PROTEIN: PS50049; TNF_157; 1.
 DR PROTEIN: PS50049; TNF_158; 1.
 DR PROTEIN: PS50049; TNF_159; 1.
 DR PROTEIN: PS50049; TNF_160; 1.
 DR PROTEIN: PS50049; TNF_161; 1.
 DR PROTEIN: PS50049; TNF_162; 1.
 DR PROTEIN: PS50049; TNF_163; 1.
 DR PROTEIN: PS50049; TNF_164; 1.
 DR PROTEIN: PS50049; TNF_165; 1.
 DR PROTEIN: PS50049; TNF_166; 1.
 DR PROTEIN: PS50049; TNF_167; 1.
 DR PROTEIN: PS50049; TNF_168; 1.
 DR PROTEIN: PS50049; TNF_169; 1.
 DR PROTEIN: PS50049; TNF_170; 1.
 DR PROTEIN: PS50049; TNF_171; 1.
 DR PROTEIN: PS50049; TNF_172; 1.
 DR PROTEIN: PS50049; TNF_173; 1.
 DR PROTEIN: PS50049; TNF_174; 1.
 DR PROTEIN: PS50049; TNF_175; 1.
 DR PROTEIN: PS50049; TNF_176; 1.
 DR PROTEIN: PS50049; TNF_177; 1.
 DR PROTEIN: PS50049; TNF_178; 1.
 DR PROTEIN: PS50049; TNF_179; 1.
 DR PROTEIN: PS50049; TNF_180; 1.
 DR PROTEIN: PS50049; TNF_181; 1.
 DR PROTEIN: PS50049; TNF_182; 1.
 DR PROTEIN: PS50049; TNF_183; 1.
 DR PROTEIN: PS50049; TNF_184; 1.
 DR PROTEIN: PS50049; TNF_185; 1.
 DR PROTEIN: PS50049; TNF_186; 1.
 DR PROTEIN: PS50049; TNF_187; 1.
 DR PROTEIN: PS50049; TNF_188; 1.
 DR PROTEIN: PS50049; TNF_189; 1.
 DR PROTEIN: PS50049; TNF_190; 1.
 DR PROTEIN: PS50049; TNF_191; 1.
 DR PROTEIN: PS50049; TNF_192; 1.
 DR PROTEIN: PS50049; TNF_193; 1.
 DR PROTEIN: PS50049; TNF_194; 1.
 DR PROTEIN: PS50049; TNF_195; 1.
 DR PROTEIN: PS50049; TNF_196; 1.
 DR PROTEIN: PS50049; TNF_197; 1.
 DR PROTEIN: PS50049; TNF_198; 1.
 DR PROTEIN: PS50049; TNF_199; 1.
 DR PROTEIN: PS50049; TNF_200; 1.
 DR PROTEIN: PS50049; TNF_201; 1.
 DR PROTEIN: PS50049; TNF_202; 1.
 DR PROTEIN: PS50049; TNF_203; 1.
 DR PROTEIN: PS50049; TNF_204; 1.
 DR PROTEIN: PS50049; TNF_205; 1.
 DR PROTEIN: PS50049; TNF_206; 1.
 DR PROTEIN: PS50049; TNF_207; 1.
 DR PROTEIN: PS50049; TNF_208; 1.
 DR PROTEIN: PS50049; TNF_209; 1.
 DR PROTEIN: PS50049; TNF_210; 1.
 DR PROTEIN: PS50049; TNF_211; 1.
 DR PROTEIN: PS50049; TNF_212; 1.
 DR PROTEIN: PS50049; TNF_213; 1.
 DR PROTEIN: PS50049; TNF_214; 1.
 DR PROTEIN: PS50049; TNF_215; 1.
 DR PROTEIN: PS50049; TNF_216; 1.
 DR PROTEIN: PS50049; TNF_217; 1.
 DR PROTEIN: PS50049; TNF_218; 1.
 DR PROTEIN: PS50049; TNF_219; 1.
 DR PROTEIN: PS50049; TNF_220; 1.
 DR PROTEIN: PS50049; TNF_221; 1.
 DR PROTEIN: PS50049; TNF_222; 1.
 DR PROTEIN: PS50049; TNF_223; 1.
 DR PROTEIN: PS50049; TNF_224; 1.
 DR PROTEIN: PS50049; TNF_225; 1.
 DR PROTEIN: PS50049; TNF_226; 1.
 DR PROTEIN: PS50049; TNF_227; 1.
 DR PROTEIN: PS50049; TNF_228; 1.
 DR PROTEIN: PS50049; TNF_229; 1.
 DR PROTEIN: PS50049; TNF_230; 1.
 DR PROTEIN: PS50049; TNF_231; 1.
 DR PROTEIN: PS50049; TNF_232; 1.
 DR PROTEIN: PS50049; TNF_233; 1.
 DR PROTEIN: PS50049; TNF_234; 1.
 DR PROTEIN: PS50049; TNF_235; 1.
 DR PROTEIN: PS50049; TNF_236; 1.
 DR PROTEIN: PS50049; TNF_237; 1.
 DR PROTEIN: PS50049; TNF_238; 1.
 DR PROTEIN: PS50049; TNF_239; 1.
 DR PROTEIN: PS50049; TNF_240; 1.
 DR PROTEIN: PS50049; TNF_241; 1.
 DR PROTEIN: PS50049; TNF_242; 1.
 DR PROTEIN: PS50049; TNF_243; 1.
 DR PROTEIN: PS50049; TNF_244; 1.
 DR PROTEIN: PS50049; TNF_245; 1.
 DR PROTEIN: PS50049; TNF_246; 1.
 DR PROTEIN: PS50049; TNF_247; 1.
 DR PROTEIN: PS50049; TNF_248; 1.
 DR PROTEIN: PS50049; TNF_249; 1.
 DR PROTEIN: PS50049; TNF_250; 1.
 DR PROTEIN: PS50049; TNF_251; 1.
 DR PROTEIN: PS50049; TNF_252; 1.
 DR PROTEIN: PS50049; TNF_253; 1.
 DR PROTEIN: PS50049; TNF_254; 1.
 DR PROTEIN: PS50049; TNF_255; 1.
 DR PROTEIN: PS50049; TNF_256; 1.
 DR PROTEIN: PS50049; TNF_257; 1.
 DR PROTEIN: PS50049; TNF_258; 1.
 DR PROTEIN: PS50049; TNF_259; 1.
 DR PROTEIN: PS50049; TNF_260; 1.
 DR PROTEIN: PS50049; TNF_261; 1.
 DR PROTEIN: PS50049; TNF_262; 1.
 DR PROTEIN: PS50049; TNF_263; 1.
 DR PROTEIN: PS50049; TNF_264; 1.
 DR PROTEIN: PS50049; TNF_265; 1.
 DR PROTEIN: PS50049; TNF_266; 1.
 DR PROTEIN: PS50049; TNF_267; 1.
 DR PROTEIN: PS50049; TNF_268; 1.
 DR PROTEIN: PS50049; TNF_269; 1.
 DR PROTEIN: PS50049; TNF_270; 1.
 DR PROTEIN: PS50049; TNF_271; 1.
 DR PROTEIN: PS50049; TNF_272; 1.
 DR PROTEIN: PS50049; TNF_273; 1.
 DR PROTEIN: PS50049; TNF_274; 1.
 DR PROTEIN: PS50049; TNF_275; 1.
 DR PROTEIN: PS50049; TNF_276; 1.
 DR PROTEIN: PS50049; TNF_277; 1.
 DR PROTEIN: PS50049; TNF_278; 1.
 DR PROTEIN: PS50049; TNF_279; 1.
 DR PROTEIN: PS50049; TNF_280; 1.
 DR PROTEIN: PS50049; TNF_281; 1.
 DR PROTEIN: PS50049; TNF_282; 1.
 DR PROTEIN: PS50049; TNF_283; 1.
 DR PROTEIN: PS50049; TNF_284; 1.
 DR PROTEIN: PS50049; TNF_285; 1.
 DR PROTEIN: PS50049; TNF_286; 1.
 DR PROTEIN: PS50049; TNF_287; 1.
 DR PROTEIN: PS50049; TNF_288; 1.
 DR PROTEIN: PS50049; TNF_289; 1.
 DR PROTEIN: PS50049; TNF_290; 1.
 DR PROTEIN: PS50049; TNF_291; 1.
 DR PROTEIN: PS50049; TNF_292; 1.
 DR PROTEIN: PS50049; TNF_293; 1.
 DR PROTEIN: PS50049; TNF_294; 1.
 DR PROTEIN: PS50049; TNF_295; 1.
 DR PROTEIN: PS50049; TNF_296; 1.
 DR PROTEIN: PS50049; TNF_297; 1.
 DR PROTEIN: PS50049; TNF_298; 1.
 DR PROTEIN: PS50049; TNF_299; 1.
 DR PROTEIN: PS50049; TNF_300; 1.
 DR PROTEIN: PS50049; TNF_301; 1.
 DR PROTEIN: PS50049; TNF_302; 1.
 DR PROTEIN: PS50049; TNF_303; 1.
 DR PROTEIN: PS50049; TNF_304; 1.
 DR PROTEIN: PS50049; TNF_305; 1.
 DR PROTEIN: PS50049; TNF_306; 1.
 DR PROTEIN: PS50049; TNF_307; 1.
 DR PROTEIN: PS50049; TNF_308; 1.
 DR PROTEIN: PS50049; TNF_309; 1.
 DR PROTEIN: PS50049; TNF_310; 1.
 DR PROTEIN: PS50049; TNF_311; 1.
 DR PROTEIN: PS50049; TNF_312; 1.
 DR PROTEIN: PS50049; TNF_313; 1.
 DR PROTEIN: PS50049; TNF_314; 1.
 DR PROTEIN: PS50049; TNF_315; 1.
 DR PROTEIN: PS50049; TNF_316; 1.
 DR PROTEIN: PS50049; TNF_317; 1.
 DR PROTEIN: PS50049; TNF_318; 1.
 DR PROTEIN: PS50049; TNF_319; 1.
 DR PROTEIN: PS50049; TNF_320; 1.
 DR PROTEIN: PS50049; TNF_321; 1.
 DR PROTEIN: PS50049; TNF_322; 1.
 DR PROTEIN: PS50049; TNF_323; 1.
 DR PROTEIN: PS50049; TNF_324; 1.
 DR PROTEIN: PS50049; TNF_325; 1.
 DR PROTEIN: PS50049; TNF_326; 1.
 DR PROTEIN: PS50049; TNF_327; 1.
 DR PROTEIN: PS50049; TNF_328; 1.
 DR PROTEIN: PS50049; TNF_329; 1.
 DR PROTEIN: PS50049; TNF_330; 1.
 DR PROTEIN: PS50049; TNF_331; 1.
 DR PROTEIN: PS50049; TNF_332; 1.
 DR PROTEIN: PS50049; TNF_333; 1.
 DR PROTEIN: PS50049; TNF_334; 1.
 DR PROTEIN: PS50049; TNF_335; 1.
 DR PROTEIN: PS50049; TNF_336; 1.
 DR PROTEIN: PS50049; TNF_337; 1.
 DR PROTEIN: PS50049; TNF_338; 1.
 DR PROTEIN: PS50049; TNF_339; 1.
 DR PROTEIN: PS50049; TNF_340; 1.
 DR PROTEIN: PS50049; TNF_341; 1.
 DR PROTEIN: PS50049; TNF_342; 1.
 DR PROTEIN: PS50049; TNF_343; 1.
 DR PROTEIN: PS50049; TNF_344; 1.
 DR PROTEIN: PS50049; TNF_345; 1.
 DR PROTEIN: PS50049; TNF_346; 1.
 DR PROTEIN: PS50049; TNF_347; 1.
 DR PROTEIN: PS50049; TNF_348; 1.
 DR PROTEIN: PS50049; TNF_349; 1.
 DR PROTEIN: PS50049; TNF_350; 1.
 DR PROTEIN: PS50049; TNF_351; 1.
 DR PROTEIN: PS50049; TNF_352; 1.
 DR PROTEIN: PS50049; TNF_353; 1.
 DR PROTEIN: PS50049; TNF_354; 1.
 DR PROTEIN: PS50049; TNF_355; 1.
 DR PROTEIN: PS50049; TNF_356; 1.
 DR PROTEIN: PS50049; TNF_357; 1.
 DR PROTEIN: PS50049; TNF_358; 1.
 DR PROTEIN: PS50049; TNF_359; 1.
 DR PROTEIN: PS50049; TNF_360; 1.
 DR PROTEIN: PS50049; TNF_361; 1.
 DR PROTEIN: PS50049; TNF_362; 1.
 DR PROTEIN: PS50049; TNF_363; 1.
 DR PROTEIN: PS50049; TNF_364; 1.
 DR PROTEIN: PS50049; TNF_365; 1.
 DR PROTEIN: PS50049; TNF_366; 1.
 DR PROTEIN: PS50049; TNF_367; 1.
 DR PROTEIN: PS50049; TNF_368; 1.
 DR PROTEIN: PS50049; TNF_369; 1.
 DR PROTEIN: PS50049; TNF_370; 1.
 DR PROTEIN: PS50049; TNF_371; 1.
 DR PROTEIN: PS50049; TNF_372; 1.
 DR PROTEIN: PS50049; TNF_373; 1.
 DR PROTEIN: PS50049; TNF_374; 1.
 DR PROTEIN: PS50049

RK MEDLINE-87184551; PubMed-3566740;
RA Young N.M.; Leon M.A.;
RT "The carbohydrate specificity of conglutinin and its homology to
RL proteins in the hepatic lectin family";
CC *Biophys. Res. Commun.* 143:645-651 (1987).
CC -*FUNCTION: CALCIUM-DEPENDENT LECTIN-LIKE PROTEIN WHICH BINDS TO A
CC YEAST CELL WALL EXTRACT & IMMUNE COMPLEXES THROUGH THE COMPLEMENT
CC N-ACETYLGLUCOSAMINE. IT IS CAPABLE OF BINDING NONREDUCING TERMINAL
CC SUBUNIT: OLIGOMERIC COMPLEX OF 4 SET OF HOMOTRIMERS.*
CC -*SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.*
CC *This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC use by non profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.ish-sib.ch/announce/>
CC or send an email to license@ish-sib.ch).*
DR EMBL; D14055; BAA03170; 1;
DR EMBL; U06860; AAB60624; 1;
DR EMBL; U06854; AAB60624; 1; JOINED.
DR EMBL; U06855; AAB60624; 1; JOINED.
DR EMBL; U06856; AAB60624; 1; JOINED.
DR EMBL; U06857; AAB60624; 1; JOINED.
DR EMBL; U06858; AAB60624; 1; JOINED.
DR EMBL; U06859; AAB60624; 1; JOINED.
DR PIR; K29416; R29416; R23140; A23140.
DR PIR; JN0450; JN0450.
DR InterPro; IPR00087; -;
DR InterPro; IPR00104; -;
DR Pfam; PF00519; lectin_c; 1.
DR PROSITE; PS00515; C-TYPE LECTIN_1; 1.
DR PROSITE; PS00041; C-TYPE LECTIN_2; 1.
KW Lectin; Hydroxylation; Glycoprotein; Mannose-binding; Membrane;
KW Collagen; Repeat; Calcium; Signal.
FT CHAIN 1 20
FT DOMAIN 21 371
FT DOMAIN 46 216
FT DOMAIN 273 371
FT MOD_RES 63 63
FT MOD_RES 87 87
FT MOD_RES 99 99
FT MOD_RES 135 135
FT MOD_RES 141 141
FT MOD_RES 159 159
FT MOD_RES 162 162
FT MOD_RES 198 198
FT SITE 201 203
FT DISULFID 275 369
FT DISULFID 347 361
FT CARBOHYD 337 337
FT CONFLICT 173 173
FT CONFLICT 210 210
FT CONFLICT 218 218
FT CONFLICT 272 272
SQ SEQUENCE 371 AA: 37994 MN: 867BA1992544B1F CRC64;
Q Query Match Similarity 35.3%; Score 878; DB 1; Length 371;
Q Best Local Similarity 49.5%; Pred. No. 3.1e 42; Mismatches 6;
Q Matches 180; Conservative 44; Mismatches 88; Indels 52; Gaps 6;
Q 1 MLLPLSMVLUQPLGNGAENKSSQRSVPNTCTLWMSPTENGLPGRDGRGREGPRG 60
Q 3 LLP-LSVLLTQPWSLGAEMTTSQKILANACHLWMSPLSGPGLPGDGDGRGCPH 61

QY 61 EKGDPGLPGPMGLSLOGPQPGPKGENGSAECPGPRGERLISGPPGLPCIPGPGAKSC 120
Db 62 EKGDPGSGPGRGAGRPGPQGPKGDGFVGPGRGPKD--TGPGRPGMPGPGAGRE 118
QY 121 PSGKQONIGPQKGPKGPKGAGPGGEWGPQGKGMOGSTGAKGSGPKGERGAGVQGAPNG 180
Db 119 PSGKQSMMPGPFTGPKGTPGPKGVAPG10--GFPGSGLAKERGAGFAGPGRG 175
QY 181 AAGPAGPACQFQGPGSGPQGPKGDRQVPGKIGKESGIPDSALRQDMEALKQKQLR 240
Db 176 VTGPGSGALQPGQFSGARCPGPGKQGDQDGPETGANGESGLAEVNLKQRTVLDGHURF 235
QY 241 EVASHYQKAALFPDGSIIRRLKIEDBRNLHEDFVFMKTORCNGERSLSLJCEELKSQ 300
Db 236 QMASQYKKAVLFPDQGAVK1-----FXTAGAVKSSDAEQL---I-CREAGK- 280
QY 301 FEGFKVDIMLNKEETKKENSFEMOKDQNPQIAHVTESEASRTTSVYQWAEKGTYMSN 360
Db 281 -----QLASPRSSAENAVENTONVRAEOKNAYLSM 310

QY 361 NIUT 364
Db 311 DIST 314

RESULT 10
ID TNF5_MOUSE
STANDARD:
AC P27548;
DT 01-AUG-1992 (Rel. 23, Created)
DT 30-MAR-2000 (Rel. 39, Last sequence update)
DR CD40 LIGAND (TNF-RELATED ACTIVATION PROTEIN) (TRAP) (T CELL ANTIGEN
DE TNFSF5 OR CD40L OR CD40L.
OS Mus musculus (Mouse)
OC Euarystora; Metacosa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-92244364; PubMed-1374165;
RA Armitage R., Fanslow W., Sato T.A., Clifford K.N., Strockbine L.,
RA Macduff B.M., Anderson D.M., Gimpel S.D., Davis-Smith T.,
RA Maliszewski C.R., Clark B.A., Smith C.A., Grabstein K.H., Cosman D.,
RT "Molecular and biological characterization of a murine ligand for
RT CD40,"
RL Nature 357:80-82(1992).
RN [2]
RP SIMILARITY TO THE TNF FAMILY.
RA MEDLINE-92310561; PubMed-1377364;
RA Farrah T., Smith C.A.;
RT "Emerging cytokine family;"
RL Nature 358:26-26(1992).
RN [3]
RP 3D-STRUCTURE MODELING OF 115-260.
RA MEDLINE-9300072; PubMed-8005800;
RA Peitsch M.C., Jongenel C.V.;
RT "A 3-D model for the CD40 ligand predicts that it is a compact trimer
RT similar to the tumor necrosis factors;"
RL Int. Immunol. 5:233-238(1993).
CC -*FUNCTION: MEDIATES B-CELL PROLIFERATION IN THE ABSENCE OF CO-
CC STIMULUS AS WELL AS IGE PRODUCTION IN THE PRESENCE OF IL-4.*
CC -*SUBUNIT: HOMODIMER.*
CC -*EXTRACELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
CC TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED ON ACTIVATED CD4+
CC CELLS.*
CC -*SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.*

OC Mammalia; Butheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
 OC Bovidae; Bovinae; Bos.
 OC NCBI_TaxID=9113;
 RN [1]
 RP SEQUENCE OF 1-242.
 RX MEDLINE=>80026026; PubMed=>488906;
 RA Kuhn K.;
 RT Fietzek P.P., Altmann H., Rauterberg J., Henkel W., Wachter E.,
 RT "The covalent structure of calf skin type III collagen. I. The amino
 acid sequence of the amino terminal region of the alpha 1(III) chain
 (positions 1-22).", Physiol. Chem. 360:809-820(1979).
 RL Hoppe-Seyler's Z. Physiol. Chem. 360:809-820(1979).
 RN [2]
 RP SEQUENCE OF 243-422.
 RX MEDLINE=>80026027; PubMed=>488907;
 RA Deves H., Fietzek P.P., Kuhn K.;
 RT "The covalent structure of calf skin type III collagen. II. The amino
 acid sequence of the cyanogen bromide peptide alpha 1(III)CB1,8,10,2
 (positions 223-402).", Physiol. Chem. 360:821-832(1979).
 RL Hoppe-Seyler's Z. Physiol. Chem. 360:821-832(1979).
 RN [3]
 RP SEQUENCE OF 423-571.
 RX MEDLINE=>80026028; PubMed=>488908;
 RA Bentz H., Fietzek P.P., Kuhn K.;
 RT "The covalent structure of calf skin type III collagen. III. The
 amino acid sequence of the cyanogen bromide peptide alpha 1(III)CB4
 (positions 403-551).", Hoppe-Seyler's Z. Physiol. Chem. 360:833-840(1979).
 RN [4]
 RP SEQUENCE OF 572-808.
 RX MEDLINE=>80026029; PubMed=>488909;
 RA Lang H., Gianielle R.W., Fietzek P.P., Kuhn K.;
 RT "The covalent structure of calf skin type III collagen. IV. The amino
 acid sequence of the cyanogen bromide peptide alpha 1(III)CB5
 (positions 552-708).", Hoppe-Seyler's Z. Physiol. Chem. 360:841-850(1979).
 RN [5]
 RP SEQUENCE OF 809-947.
 RX MEDLINE=>80026030; PubMed=>488910;
 RA Dexes H., Fietzek P.P., Kuhn K.;
 RT "The covalent structure of calf skin type III collagen. V. The amino
 acid sequence of the cyanogen bromide peptide alpha 1(III)CB9A
 (position 789-927).", Hoppe-Seyler's Z. Physiol. Chem. 360:851-860(1979).
 RN [6]
 RP SEQUENCE OF 948-1049.
 RX MEDLINE=>80026031; PubMed=>488911;
 RA Allmann H., Fietzek P.P., Gianielle R.W., Kuhn K.;
 RT "The covalent structure of calf skin type III collagen. VI. The amino
 acid sequence of the carboxy-terminal cyanogen bromide peptide alpha
 1(III)CB9B (positions 928-1028).", Hoppe-Seyler's Z. Physiol. Chem. 360:861-868(1979).
 RT -I- FUNCTION: COLLAGEN TYPE III OCCURS IN MOST SOFT CONNECTIVE TISSUES
 CC ALONG WITH TYPE I COLLAGEN.
 CC -!- SUBUNIT: TRIMERS OF IDENTICAL ALPHA 1(III) CHAINS. THE CHAINS ARE
 CC LINKED TO EACH OTHER BY INTERCHAIN DISULFIDE BONDS. TRIMERS ARE
 CC ALSO CROSS-LINKED VIA HYDROXYLISINES.
 CC -!- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPETIDE REPEATING
 CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
 CC PIR: A03862; CGB075.
 DR InterPro; IPR000877; -
 DR InterPro; IPR001007; -
 DR Pfam; PF01391; Collagen: 17;
 DR PROSITE; PS01208; VWFC: PARTIAL.
 KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
 KW Glycoprotein; Collagen.
 DR DOMAIN 1 14 NONHELCICAL REGION (N-TERMINAL).
 DR DOMAIN 15 1040 TRIPLE-HELICAL REGION.
 DR DOMAIN 1041 1049 NONHELCICAL REGION (C-TERMINAL).
 DR MOD_BRS 95 95 HYDROXYLATION.
 DR MOD_RES 107 107 HYDROXYLATION.
 DR MOD_RES 119 119 HYDROXYLATION.
 DR MOD_RES 938 938 HYDROXYLATION.

FT MOD_RES 950 950 HYDROXYLATION.
 FT CARBOHD 107 107 CROSS-LINK SITE.
 FT CARBOHTD 950 950 CROSS-LINK SITE.
 FT DISULFID 1040 1040 INTERCHAIN.
 FT DISULFID 1041 1041 INTERCHAIN.
 SQ 1049 AA; 93651 MM; 8EBC33DC66EC9A3 CRC64;

Query Match 19.5%; Score 484.5; DB 1; Length 1049;
 Best Local Similarity 41.2%; Pred. No. 5.2e-20;
 Matches 113; Conservative 13; Mismatches 71; Index 77; Gaps 9;

QY 14 PLGNIGAENKLSLSQRSVNTCTLVMCSPTENGLPDRGRGRE----GPROBKGDGL 67
 Db 427 PKGNDGARGK-N-GERGCPG----GPGPOGPAGKNGENGPQGPGPQGPGSDKGPG 478
 QY 68 PGPMGLSLQGPQGPVPGENGSAECPGRGERL----SCPPGLP 112
 Db 479 PGPGLOGLPGTS--GPPGENGKPGCPGKGEAGAPGIPGIPGGKGDGAPGPG 535
 QY 113 PEGAKGKGPSKGQGNIGPQKPGPKGEA----GPKGEVAPGNGQS 154
 Db 536 PEGPGKGPQPGPKGPQEGKGAAQDPPGPGSAGTPGQGMPGERGGPGPQPKGEPOSSGV 595
 QY 155 TGAKGSTGP-----KGERGAPGV----QGAPGNAGLAGPAGPAGP 190
 Db 596 DQAPGKGDPGRPTGPIGPPCPAGPAGPQGDKEGGAPGVPIAGPQGGKGERGEGQPGPAGF 655
 QY 191 OGAPGSRGPGLKGDRGVRQDRGKGEGLP 224
 Db 656 PGAPGONGSGEPAKGERGAPGE--KGEGGPGAA 686

RESULT 13
 CALL_CANFA
 ID CALL_CANFA STANDARD; PRT; 1460 AA.
 AC Q9X577;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DE COLLAGEN ALPHA 1(I) CHAIN PRECURSOR.
 GN COL1A1
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
 OC Mammalia; Carnivora; Fissipedia; Canidae; Canis.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-SKIN
 RA Campbell B.G., Woottton J.A.M., McLeod J.N., Minor R.R.;
 CC Sequence of normal canine COL1A1 cDNA.;
 RT Submitted (MAY-1999) to the EMBL/Genbank/DBJ databases.
 CC -!- FUNCTION: TYPE I COLLAGEN IS A MEMBER OF GROUP I COLLAGEN
 CC (FIBRILLAR FORMING COLLAGEN).
 CC -!- SUBUNIT: TRIMERS OF ONE ALPHA 2(I) AND TWO ALPHA 1(I) CHAINS.
 CC -!- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPETIDE REPEATING
 CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
 CC -!- SIMILARITY: CONTAINS 1 VWFC DOMAIN.
 CC
 This SWISS PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>).
 DR EMBL; AF153062; AAD4619.1; -
 DR InterPro; IPR000875; -
 DR InterPro; IPR001007; -
 DR Pfam; PF0410; COL1I; 1
 DR PROSITE; PS01208; VWFC; 1
 KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;

KW GLYCOPROTEIN; Collagen; Signal.
 PT SIGNAL 1 22 BY SIMILARITY.
 PT PROPEP 23 157 AMINO-TERMINAL PROPEPTIDE.
 PT CHAIN 158 1214 COLLAGEN ALPHA 1(I) CHAIN.
 PT PROPEP 1215 1460 CARBOXYL-TERMINAL PROPEPTIDE.
 PT DOMAIN 134 92 VWL.
 PT DOMAIN 158 174 NONHICAL REGION (N-TERMINAL).
 PT DOMAIN 175 1188 TRIPLE-HELICAL REGION.
 PT DOMAIN 1189 1214 NONHICAL REGION (C-TERMINAL).
 PT SITE 741 743 CELL ATTACHMENT SITE (POTENTIAL).
 PT SITE 1089 1091 CELL ATTACHMENT SITE (POTENTIAL).
 PT CARBOHYD 1361 1361 N-LINKED (GLCNAC) (POTENTIAL).
 SQ SEQUENCE 1460 AA; 138762 MW; 58E3674D2B570697 CRC64;

Query Match 19.4%; Score 482.5; DB 1: Length 1460;
 Best Local Similarity 42.0%; Pred. No. 9.4e-20; Matches 113; Conservative 14; Mismatches 71; Indels 71; Gaps 8;

Oy 14 PLGNLGAEMKSLRSQSPVNPCTLVNCSPPEENGFLGRDGRGPRGRKGDKPGLPGMNL 73
 Db 848 PIGNVGAPGPKGKGSGAG----PFCATGPFGAAGRVGPPGPGSGHAGPPGPGAGK 899

Oy 74 SGQ-----GPTCPVPGKGENGAGEPOPQKGEGFLSGPGL----- 109
 Db 900 EGGKGARGETGPAGRPGEVGPQGPQGPAGKSGPAGDQGPGQIAGORGVYGL 959

Oy 110 -----PGIOPA----GKEGPGSKGKDNIGQKGPKGPKGEAPKGE--VCAQMGKGS 154
 Db 960 PGQERGERGPGLCPGSGKQCPGSGTSGERGPQGPMPGLAGPQSGSGRECGSPGAGS 1019

Oy 155 TGAKGSTGPKRGER-----GAPVGQGAPGNGAA-----GAPGACQGARGS 196
 Db 1020 PGRCGSPGPKGDRGETGPAGPQGAPGAGPAGPQGPVGA 1079

Oy 197 RGPGL-----KGDRGPQGDKGPKGKGESC 219
 Db 1080 RGPAGQGPQGDKGETGEQOGDRGKKGHRG 1108

RESULT 14
 CALL_HUMAN STANDARD: PRT: 1464 AA.
 ID P02452; 015176; 014037;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE COLLAGEN ALPHA 1(I) CHAIN PRECURSOR.
 GN COL1A1.
 OS Homo sapiens (Human).
 OC Chordata; Craniata; vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homino;
 OC NCBI_TaxID=9606;
 RN [1] SEQUENCE OF 1-472 FROM N.A.
 RX MEDLINE-B89021644; PubMed=3118743;
 RA Tramp G., Kulvaniemi H., Stacey A., Shikata H., Baldwin C.T.,
 RA Jaenisch R., Prockop D.J.,
 RT "Structure of a full-length cDNA clone for the prepro alpha 1(I)
 chain of human type I procollagen.",
 RL Biochem. J. 253:919-922(1988).
 RN [2] SEQUENCE OF 1-181 FROM N.A.
 RX MEDLINE-B84270697; PubMed=642220;
 RA Chu M.-L., de Wet W.J., Bernard M.P., Ding J.-F., Morabito M.,
 RA Myers J., Williams C., Ramirez F.;
 RT "Human pro alpha 1(I) collagen gene structure reveals evolutionary
 conservation of a pattern of introns and exons.",
 RL Nature 310:337-340(1984).
 RN [3] SEQUENCE OF 162-301.
 RX MEDLINE-71038625; PubMed=5529814;

RA Click E.M., Bernstein P.;
 RT "Isolation and characterization of the cyanogen bromide peptides from
 the alpha 1 and alpha 2 chains of human skin collagen.",
 RT Biochemistry 9:4699-4706(1970).
 RN [4] SEQUENCE OF 263-268.
 RC TISSUE=Skin;
 RX MEDLINE-71001508; PubMed=4319110;
 RA Morgan P.H., Jacobs H.G., Segrest J.P., Cunningham L.W.;
 RT "A comparative study of glycopeptides derived from selected
 vertebrate collagens. A possible role of the carbohydrate in fibril
 formation.", J. Biol. Chem. 245:5042-5048(1970).
 RN [5] SEQUENCE OF 425-164 FROM N.A.
 RX MEDLINE-B84080385; PubMed=6680127;
 RA Bernard M.P., Chu M.-L., Myers J.C., Ramirez F., Eikenberry E.F.,
 RA Prockop D.J.;
 RT "Nucleotide sequences of complementary deoxyribonucleic acids for the
 pro alpha 1 chain of human type I Procollagen. Statistical evaluation
 of structures that are conserved during evolution.",
 RT Biochemistry 22:5213-5223(1983).
 RN [6] SEQUENCE OF 1229-1454 FROM N.A.
 RC TISSUE=Bone;
 RX MEDLINE-B88124108; PubMed=3340531;
 RA Maeckle J.K., Raassina M., Virta A., Vuorio E.;
 RT "Human pro alpha 1(I) collagen: cDNA sequence for the C-propeptide
 domain.", Nucleic Acids Res. 16:349-349(1988).
 RN [7] SEQUENCE OF 1-34 FROM N.A.
 RX MEDLINE-B88097389; PubMed=3480516;
 RA Borstein P., McKay J., Morishima J.K., Devarayalu S., Gelinas R.E.;
 RT "Regulatory elements in the first intron contribute to
 transcriptional control of the human alpha 1(I) collagen gene.",
 RL Proc. Natl. Acad. Sci. U.S.A. 84:8869-8873(1987).
 RN [8] SEQUENCE OF 1-34 FROM N.A.
 RX MEDLINE-B85130370; PubMed=2857713;
 RA Chu M.-L., de Wet W.J., Bernard M.P., Ramirez F.;
 RT "Fine structural analysis of the human pro-alpha 1 (I) collagen gene.
 RT Promoter structure, Alu repeats, and polymorphic transcripts.",
 RL J. Biol. Chem. 260:2315-2320(1985).
 RN [9] SEQUENCE OF 1-44 FROM N.A.
 RX MEDLINE-B88033098; PubMed=2822714;
 RA ROSSOUW C.M.S., Vergerre W.P., du Plooy S.J., Bernard M.P., Ramirez F.,
 RA de Wet W.J.;
 RT "DNA sequences in the first intron of the human pro-alpha 1(I)
 collagen gene enhance transcription.",
 RL J. Biol. Chem. 262:15151-15157(1987).
 RN [10] REVIEW ON VARIANTS.
 RX MEDLINE-91180577; PubMed=2010058;
 RA Kulvaniemi H., Tramp G., Prockop D.J.;
 RT "Mutations in fibrillar collagens (types I, II, III, and XI), fibrillar
 collagen genes: causes of rare and some common diseases
 in humans.",
 RL FASEB J. 5:2052-2060(1991).
 RN [11] REVIEW ON VARIANTS.
 RX MEDLINE-97255959; PubMed=9101290;
 RA Kulvaniemi H., Tramp G., Prockop D.J.;
 RT "Mutations in fibrillar collagens (types I, II, III, and XI), fibrillar
 collagen genes: causes of rare and some common diseases
 in humans.",
 RL Hum. Mutat. 9:300-315(1997).
 RN [12] REVIEW ON OI VARIANTS.
 RX MEDLINE-91374476; PubMed=1895312;
 RA Byers P.H., Wallis G.A., Willing M.C.;
 RT "Osteogenesis imperfecta: translation of mutation to phenotype.",
 RL J. Med. Genet. 28:433-442(1991).

[13] REVIEW ON OI VARIANTS.

RP MEDLINE=97169389; PubMed=9016532;

RA Dalgleish R.;

RT "The human type I collagen mutation database.";

RL Nucleic Acids Res. 25:811-87 (1997).

RN [14] VARIANT OI-II CYS-116.

RP MEDLINE=82287390; PubMed=3016737;

RA Conn D.H., Byers P.H., Steinmann B., Gelineas R.E.,

RT "Lethal perinatal osteogenesis imperfecta resulting from a single nucleotide change in one human pro alpha 1(I) collagen allele.";

RL Proc. Natl. Acad. Sci. U.S.A. 83:6045-6047 (1986).

RN [15] VARIANT OI-II ARG-569.

RP MEDLINE=8722295; PubMed=3108247;

RA Bateman J.F., Chan D., Walkers I.D., Rogers J.G., Cole W.G.,

RT "Lethal perinatal osteogenesis imperfecta due to the substitution of arginine for glycine at residue 391 of the alpha 1(I) chain of type I collagen.";

RL J. Biol. Chem. 262:7021-7027 (1987).

RN [16] VARIANT OI-II CYS-926.

RP MEDLINE=88033031; PubMed=3667599;

RA Vogel B.E., Minor R.R., Freund M., Prockop D.J.,

RT "A point mutation in a type I procollagen gene converts glycine 748 of the alpha 1 chain to cysteine and destabilizes the triple helix in a lethal variant of osteogenesis imperfecta.";

RL J. Biol. Chem. 262:14737-14744 (1987).

RN [17] VARIANT OI-II ARG-842.

RP MEDLINE=88298028; PubMed=3403550;

RA Bateman J.F., Lamande S.R., Dahl H.H., Chan D., Cole W.G.,

RT "Substitution of arginine for glycine 664 in the collagen alpha 1(I) chain in lethal perinatal osteogenesis imperfecta. Demonstration of the peptide defect by in vitro expression of the mutant cDNA.";

RL J. Biol. Chem. 263:11627-11630 (1988).

RN [18] VARIANT OI-II CYS-1195.

RP MEDLINE=89218620; PubMed=344312;

RA Labhard M.E., Wirtz M.K., Pope F.M., Nicholls A.C., Hollister D.W.,

RT "A cysteine for glycine substitution at position 1017 in an alpha 1(I) chain of type I collagen in a patient with mild dominantly inherited osteogenesis imperfecta.";

RL Mol. Biol. Med. 5:197-207 (1988).

RN [19] VARIANT OI-II VAL-34.

RP MEDLINE=89255493; PubMed=2470760;

RA Patterson E., Smiley E., Bonadio J.,

RT "RNA sequence analysis of a perinatal lethal osteogenesis imperfecta mutation";

RL J. Biol. Chem. 264:10083-10087 (1989).

RN [20] VARIANT OI-IV SER-1010.

RP MEDLINE=8930591; PubMed=2745420;

RA Martini J.C., Grange D.K., Gottesman G.S., Lewis M.B., Koeplin D.A.,

RT "Osteogenesis imperfecta type IV. Detection of a point mutation in one alpha 1(I) collagen allele (COL1A1) by RNA/RNA hybrid analysis.";

RL J. Biol. Chem. 264:11893-11900 (1989).

RN [21] VARIANT OI-II ALA-1106; VAL-1151; ARG-1154 AND VAL-1184.

RP MEDLINE=89380165; PubMed=2777764;

RA Lamande S.R., Dahl H.-H.M., Cole W.G., Bateman J.F.,

RT "Characterization of point mutations in the collagen COL1A1 and COL1A2 genes causing lethal perinatal osteogenesis imperfecta.";

RL J. Biol. Chem. 264:15809-15812 (1989).

RN [22] VARIANT OI SER-1022.

RP MEDLINE=90062668; PubMed=2511192;

RA Pack M., Constantino C.D., Kalla K., Nielsen K.B., Prockop D.J.,

RT "Substitution of serine for alpha 1(I)-glycine 841 in a severe variant of osteogenesis imperfecta minimally destabilizes the triple helix of type I procollagen. The effects of glycine substitutions on

RT thermal stability are either position of amino acid specific.";

RL J. Biol. Chem. 264:19694-19699 (1989).

RN [23] VARIANT OI-II CYS-1192.

RP MEDLINE=89109573; PubMed=2913053;

RA Constantinou C.D., Nielsen K.B., Prockop D.J.,

RT "A lethal variant of osteogenesis imperfecta has a single base mutation that substitutes cysteine for glycine 904 of the alpha 1(I) chain of type I procollagen. The asymptomatic mother has an unidentified mutation producing an overmodified and unstable type I procollagen.";

RL J. Clin. Invest. 83:574-584 (1989).

RN [24] VARIANT OI CYS-27; CYS-704 AND CYS-896.

RP MEDLINE=9000313; PubMed=2794057;

RA Stamman B.J., Eyre D., Charbonneau H., Harrylock M., Weiss M.A.,

RT Weiss L., Graham J.M., Byers P.H.,

RL "Osteogenesis imperfecta. The position of substitution for glycine by cysteine in the triple helical domain of the pro alpha 1(I) chains of type I collagen determines the clinical phenotype.";

RN J. Clin. Invest. 84:1206-1214 (1989).

RP VARIANT OI-II CYS-422.

Query Match Best Local Similarity 19.4%; score 481.5; DB 1; Length 1464; Matches .112; Conservative 9; Mismatches 77; Index 71; Gaps 7;

QY 14 PLNIGAEMMSLQSRSVPNCTLVMCSPTENGLPGRGRCRDRGREGPKEGKPGIPLGKGML 73

Db 852 PIGNYVGAPKGANGSAG-----PPGATCFFPGAGAAGRVPPGSGNSNAGPPGIPGPACK 903

QY 74 SGLOQPTGPGPKENGSAECP-----IGERGLSGPGLPGLPGP----- 115

Db 904 EGKGPGRGETGPAGRPGPVGPGPPGPAGERGPAGSPGAGPAGPAPSPGPAGIAGRCVGL 963

QY 116 -----AGKGPSPKGKGNNGPQGKPGPKGE-----VGA 148

Db 964 PGQRGERGFPGLGPSPCEPGKPGPSPGASGERGPGPMPGMPGPGLAGPPGSGSGPAGREGS 1023

QY 149 PGHQGSGTAK---GSGPKGRGARGVGQVAGPGNAGA-----GPAQPGQAGPS 196

Db 1024 PGDRGSPGAKDGRGEGTPAGPGAGPAGPAGPAGPVGPGACKSGDGRGETGPAGPGVPGAGA 1083

QY 197 RPPGL-----KGDRGVPSDRGIGESE 219

Db 1084 RGPAGPQGPDRKGEGEOGCDRGKGHRG 1112

RESULT 15

CALL CHICK ID CALL CHICK STANDARD; PRTR; 1453 AA.

AC P02457; DT 21-JUL-1986 (Rel. 01, Created)

RT 01-OCT-1989 (Rel. 12, Last sequence update)

RL 15-JUL-1999 (Rel. 38, Last annotation update)

DE COLLAGEN ALPHA 1(I) CHAIN PRECURSOR.

GN COL1A1.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Archosia; Ave; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.

OX NCBI_TaxID=9031; RN [1]

RP SEQUENCE OF 1-153 FROM N.A.

RX MEDLINE=88056316; PubMed=3678834;

RA Finer M.H., Boedeker H., Doty P.,

RT "Construction and characterization of cDNA clones encoding the 5' end of the chicken pro alpha 1(I) collagen mRNA.";

RL Gene 56:71-78 (1987).

RN [2]

RP SEQUENCE OF 1-144 FROM N.A.

RX MEDLINE=88007542; PubMed=24220966;

THIS PAGE BLANK (version)

GenCore version 4.5
 Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using SW model

Run on: July 3, 2001, 16:14:38 ; Search time 25.5 seconds
 (without alignments)
 2443.752 Million cell updates/sec

Title: US-09-454-223-2
 Perfect score: 2487
 Sequence: 1 MLPFLSLMLVLLVPLGNLGA.....TDPSQVSHGTFNSFGLLKL 471.

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SPTRMBL_16:*

1: sp_archea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp_rhodent:*

12: sp_unclassified:*

13: sp_vertebrate:*

14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	1039.5	41.8	378	6 Q9N1X4
2	836	33.6	260	11 Q922V2
3	827	33.3	260	11 Q9R254
4	697	28.0	335	6 Q97748
5	499.5	20.1	890	5 Q77087
6	488.5	19.6	888	13 Q90796
7	48.8	19.6	1835	13 Q9IAU4
8	482.5	19.4	1445	13 Q93251
9	481	19.3	1431	11 Q9JMH4
10	480.5	19.3	1461	4 Q76045
11	480.5	19.3	1838	11 Q88207
12	480	19.3	1767	4 Q90174
13	480	19.3	1806	4 Q9U75
14	480	19.3	1818	4 Q9U76
15	479.5	19.3	1838	4 Q15094
16	479.5	19.3	1840	11 Q60467
17	479	19.3	1840	11 Q93103
18	478.5	19.2	1453	11 Q63079
19	478	19.2	1736	4 Q9UTP9

ALIGMENTS

RESULT	1	Q9N1X4	PRELIMINARY;	PRY;	378 AA.
ID	Q9N1X4;				
AC	Q9N1X4;				
DT	01-OCT-2000	(TREMBlrel. 15, Last sequence update)			
DT	01-MAR-2001	(TREMBlrel. 16, Last annotation update)			
DE	LUNG SURFACTANT PROTEIN D PRECURSOR.				
GN	SFPD.				
OS	Sus scrofa (PIG).				
OC	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Buteleostomi; OX				
NCBI_TAXID	9823;				
RN	[1]				
RP	SEQUENCE FROM N.A. MEDLINE-2010598; PUBMED=10640760;				
RX	EMBL; AF132496; ARF2215.2; ..				
RA	Lawson P.R.; Portcine Lung Surfactant Protein D (SP-D): cDNA cloning, chromosomal localisation and tissue distribution.*; van Eijk M., Haagman H.P., Skinner T., Archibald A.L., Reid K.B.M., EMBL; AF132496; ARF2215.2; ..				
RA	InterPro: IPR00087; ..				
RT	InterPro: IPR001304; ..				
RT	DR Pfam: PF00059; Lectin_C; 1.				
RL	DR Pfam; PF01391; Collagen; 3.				
DR	PROSITE; PS00615; C_TYPE-LECTIN_1; 1.				
DR	PROSITE; PS50041; C_TYPE-LECTIN_2; 1.				
DR	SMART; SM00044; CLECT; 1.				
KW	Signal.				
FT	SIGNAL	1	20	POTENTIAL.	
FT	CHAIN	21	378	LUNG SURFACTANT PROTEIN D.	
SQ	SEQUENCE	378 AA;	37986 MN;	3504E8C1E56C341D CRC64;	

Query Match 41.8%; Score 1039.5; DB 6; Length 378; Best Local Similarity 55.8%; Pred. No. 3.5e-73; Matches 218; Conservative 38; Mismatches 82; Index 53; Gaps 8;

QY 1 MLPFLSLMLVLLVPLGNLGA.....TDPSQVSHGTFNSFGLLKL 471

Db 3 LLPL-SLVLLTLLTOPRSLGAEMKTYQSRAVANACALVNCSPSPENGIPGRDGDRGREGPRG 60

Q9uip9 homo sapien

QY 61 EKDPGLPGLCPMGSLGLOOPTGPVGPKGENGSAGERGPCKGERGLSGPPGLPGLPAGKEG 120
 Db 62 EKDPGLPGLPAGVGRAGMGLAFLPGVKGDRGNGSTGERGAKGDGPGCPGPRPGTGPAGKEG 121
 QY 121 PSCKQKGNGPQPKPGKGEACPKGEKGPMQGSGAKSGRGERAPGVGAPSNAG 180
 Db 122 PSQCGQGNIGPQPGTGPQKGKGERGPKGEVGAQGALGMQGSGTARGPAGLKGGERGAPGDSAG 181
 QY 181 AAGPAGATGCPGCPGSGARGPGLKGDRGPGTGAQVNLQAGQMASPSET 344
 Db 182 AAGPAGATGCPGCPGSGARGPGLKGDRGPGTGAQVNLQAGQMASPSET 344
 QY 241 EVAFSHQKALFPDGHRRDKIEDERNLHEDFVPMKTIQ-----RCNT 284
 Db 242 QKAFSQYKKELFNPGNGVGEKI-----FKTGFEKFQDAAQVOTQAGQMASPSET 295
 QY 285 GERSLSLNLCEELISOSGFSVQDILMKNEETKENSEMOKDQDNQIAAHVISASKT 344
 Db 296 SNEALSQUTAQNAKAFL-----LSMTDIRKTEGNNFTYPGE-PLWYAN----- 336
 QY 345 TSVIQLWA-----ERGYYMSNIVL-ENG 369
 Db 337 -----WAEPEPNNGGSSGAENCVIFPNK 362

RESULT 2
 ID Q9ZV2
 ID Q9ZV2; PRELIMINARY; PRT; 260 AA.
 AC 09ZV2;
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathl; Muridae; Murinae; Rattus.
 OC Mammalia; Eutheria; Rodentia; Sciurognathl; Muridae; Murinae; Rattus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE_DAWLEY;
 RA Hallert K.M., Oaks M.K.;
 RT "Nucleotide Sequence of the Rat CD40 Ligand.",
 RL Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
 DR EMBL: AF116582; AAD22460; 1;
 DR HSSP: P27548; ICDR:
 DR InterPro: IPR00478; -.
 DR Pfam: PF00229; TNF; 1.
 DR PROSITE: PS00251; TNF; 1.
 DR SMART: SM0027; TNF; 2; 1.
 DR SEQUENCE: 260 AA; 29259 MW; B3D3757DB60DB73A CRC64;

Query Match 33.3%; Score 827; DB 11; Length 260;
 Best Local Similarity 48.8%; Pred. No. 7.7e-57;
 Matches 171; Conservative 28; Mismatches 53; Indels 12; Gaps 4;
 RT "Cloning and sequencing of rat cd40 ligand.",
 RL Submitted (JUL-1997) to the EMBL/GenBank/DDBJ databases.
 DR EMBL: AF013985; AAD09523; 1; -.
 DR HSSP: P27548; ICDR:
 DR InterPro: IPR00478; -.
 DR Pfam: PF00229; TNF; 1.
 DR PROSITE: PS00251; TNF; 1.
 DR SMART: SM0027; TNF; 1.
 DR SEQUENCE: 260 AA; 29282 MW; 15D21F5200F6DEBB CRC64;

Query Match 33.6%; Score 836; DB 11; Length 260;
 Best Local Similarity 65.5%; Pred. No. 1.5e-57;
 Matches 173; Conservative 27; Mismatches 52; Indels 12; Gaps 4;
 QY 209 PGDRGKGEGLPSAALRQMEALKGKQRL-EVAFSHQKALFPDGHRRDKIEDER 267
 Db 8 PSRPSV-ATGLPASMKIFMVLITVFLTOMIGSPVAVYL-----HRLDKVEEA 57
 QY 268 NLHEDEVFMKTIQRCNTGERSLSLNLCEELISOSGFSVQDILMKNEETKENSEMOKD 327
 Db 58 SLHEDEVFMKTKRCKNGCGSLSLNLCEEMOKDPEVDKDSLNLKE-KIEKSEMOGRD 116
 QY 328 ONPQTAHAHTVSEASSKTTSVIQLWAEKGYTMNSNIVLLENKQLTVKROGLYYAQTF 387
 Db 117 EDPOIAAHVSEANSNAASVLOKAKGYTMNSNIVLLENGLRPLVKGSLYVYQTIF 176
 QY 388 CSNREASSGAPFASLCLKSPRFERTILRANTHSAKCGQSIHLAGVFLQPGASV 447
 Db 177 CSNREPLSORPFIIVSLWIKPSSGSRFILLRANTHSKLCQDQSIHLAGVFLQAGASV 236
 QY 448 FVNVTDPQSIVHTGFGFLKL 471
 Db 237 FVNVTESASQVINGIGFSSIGLKL 260

RESULT 3
 ID Q9R354
 ID Q9R354; PRELIMINARY; PRT; 260 AA.
 AC 09R354;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE CD40 LIGAND
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathl; Muridae; Murinae; Rattus.
 OC NCBI_TAXID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE_DAWLEY;
 RA Hallert K.M., Oaks M.K.;
 RT "Nucleotide Sequence of the Rat CD40 Ligand.",
 RL Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
 DR EMBL: AF116582; AAD22460; 1;
 DR HSSP: P27548; ICDR:
 DR InterPro: IPR00478; -.
 DR Pfam: PF00229; TNF; 1.
 DR PROSITE: PS00251; TNF; 1.
 DR SMART: SM0027; TNF; 1.
 DR SEQUENCE: 260 AA; 29259 MW; B3D3757DB60DB73A CRC64;

Query Match 33.3%; Score 827; DB 11; Length 260;
 Best Local Similarity 48.8%; Pred. No. 7.7e-57;
 Matches 171; Conservative 28; Mismatches 53; Indels 12; Gaps 4;
 RT "Cloning and sequencing of rat cd40 ligand.",
 RL Submitted (JUL-1997) to the EMBL/GenBank/DDBJ databases.
 DR EMBL: AF013985; AAD09523; 1; -.
 DR HSSP: P27548; ICDR:
 DR InterPro: IPR00478; -.
 DR Pfam: PF00229; TNF; 1.
 DR PROSITE: PS00251; TNF; 1.
 DR SMART: SM0027; TNF; 1.
 DR SEQUENCE: 260 AA; 29282 MW; 15D21F5200F6DEBB CRC64;

Query Match 33.6%; Score 836; DB 11; Length 260;
 Best Local Similarity 65.5%; Pred. No. 1.5e-57;
 Matches 173; Conservative 27; Mismatches 52; Indels 12; Gaps 4;
 QY 209 PGDRGKGEGLPSAALRQMEALKGKQRL-EVAFSHQKALFPDGHRRDKIEDER 267
 Db 8 PSRPSV-ATGLPASMKIFMVLITVFLTOMIGSPVAVYL-----HRLDKVEEA 57
 QY 268 NLHEDEVFMKTIQRCNTGERSLSLNLCEELISOSGFSVQDILMKNEETKENSEMOKD 327
 Db 58 SLHEDEVFMKTKRCKNGCGSLSLNLCEEMOKDPEVDKDSLNLKE-KIEKSEMOGRD 116
 QY 328 ONPQTAHAHTVSEASSKTTSVIQLWAEKGYTMNSNIVLLENKQLTVKROGLYYAQTF 387
 Db 117 EDPOIAAHVSEANSNAASVLOKAKGYTMNSNIVLLENGLRPLVKGSLYVYQTIF 176
 QY 388 CSNREASSGAPFASLCLKSPRFERTILRANTHSAKCGQSIHLAGVFLQPGASV 447
 Db 177 CSNREPLSORPFIIVSLWIKPSSGSRFILLRANTHSKLCQDQSIHLAGVFLQAGASV 236
 QY 448 FVNVTDPQSIVHTGFGFLKL 471
 Db 237 FVNVTESASQVINGIGFSSIGLKL 260

RESULT 4
 ID 097748
 ID 097748; PRELIMINARY; PRT; 335 AA.
 AC 097748;
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE CONGUTININ PRECURSOR.
 OS Bos taurus (Bovine).
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

	RESULT	5
RC	077087	PRELIMINARY; PRT; 890 AA.
AC	077087;	
DR	01-NOV-1998 (TREMBLrel. 08, Created)	
DT	01-NOV-1998 (TREMBLrel. 12, Last sequence update)	
DT	01-MAR-2001 (TREMBLrel. 16, Last annotation update)	
DE	FIBRILLAR COLLAGEN CHAIN FAPI ALPHA.	
OS	Alvinella pompejana	
OC	Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Canalipalpata;	
OC	Terebellida; Alvinellidae; Alvinella.	
OX	NCBI_TAXID=6376;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=LIVER;	
RX	MEDLINE=3321361; PubMed=8460993;	
RA	Suzuki Y., Yin Y., Makino M., Kurimura T., Wakamiya N.;	
RT	Cloning and sequencing of a cDNA coding for bovine conglutinin.;	
RL	Biochem. Biophys. Res. Commun. 191:335-342(1993).	
RL	[2]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=LIVER;	
RX	MEDLINE=3277552; PubMed=7684896;	
RA	Liu J., Lauritsen S., Thiel S., Janssenius J., Reid K.;	
RT	"The cDNA cloning of conglutinin and identification of liver as a primary site of synthesis of conglutinin in members of the Bovidae.;"	
RL	Biochem. J. 292:157-162(1993).	
RL	[3]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=LIVER;	
RX	MEDLINE=91128104; PubMed=8297370;	
RA	Kawasaki N., Itoh N., Kawasaki T.;	
RT	"Gene organization and 5'-flanking region sequence of conglutinin: a C-type mammalian lectin containing a collagen-like domain.;"	
RL	Biochem. Biophys. Res. Commun. 198:597-604(1994).	
DR	EMBL; D25302; BAA04983..1; -.	
DR	EMBL; D25296; BAA04983..1; JOINED.	
DR	EMBL; D25297; BAA04983..1; JOINED.	
DR	EMBL; D25299; BAA04983..1; JOINED.	
DR	EMBL; D25300; BAA04983..1; JOINED.	
DR	EMBL; D25301; BAA04983..1; JOINED.	
DR	EMBL; P35247; 1B08.	
DR	InterPro; IPR000087; -.	
DR	InterPro; IPR001304; -.	
DR	Pfam; PF00059; Lectin_c; 1.	
DR	Pfam; PF01391; Collagen; 10.	
DR	PROSITE; PS00615; C-TYPE-LECTIN_1; 1.	
DR	PROSITE; PS00034; CLECT_1; 1.	
DR	SMART; SM0034; CLECT_1; 1.	
FT	CONFLICT 236 236 E -> V (IN REF 2).	
FT	SMART; SM0034; CLECT_1; 1.	
SEQUENCE	335 AA; 34702 MW; 39b3a30bc76c134c CRC64;	
Query Match	20.1%	Score 499.5; DB 5; Length 890;
Best Local Similarity	47.5%	Pred. No. 1.4e-46;
Matches	115;	Conservative 40; Mismatches 84; Gaps 6;
Indels	43;	Gaps 6;
Qy	14 PLGNIG-----ABMKSLRSORSVNPNTCTLMVCS-----TENGUQDRGDGRGSPRGEKG 63	
Db	320 PAGNTGPGPGPAGCIKGL--RGEFG-----KFGRADGPFGTPGMDGTGKGERGEDG 370	
Qy	64 DPGCLPGPMGLSGLQ-----GPTCPWPGPKENGNSAGEP--GPKGGERGLSG 105	
Db	371 PPGLPGTPGPQCBERGFVGLPGGRGEQPGFPGPQGPQGPDTGPGRERGSPG 430	
Qy	106 PPGLPGIPGPAGKEGPSGKQGNIGQKGP--GPKGAEQPKGEVQGAPGMQSGTAKGKSTG 162	
Db	431 PSGAGPAGPAGPSPPPGKPLAGRPGDKGPQPPPGLAGPMQGLRGPPG 490	
Qy	163 PKGEGPAGPVGAPCAGNAGPAGRA--GPOQAGRSRGPGLKGDRGVPGDRGKIGRSQG 219	
Db	491 PGGEGERGERGPAGCVRGVPGVPGPQGAGPQGQGERGAAAGPKGDKGPGMPLQGMQG 550	
Qy	220 LP 221	
Db	551 TP 552	
RESULT	6	
Q90796	PRELIMINARY; PRT; 888 AA.	
AC	090796;	
DT	01-NOV-1996 (TREMBLrel. 01, Created)	
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)	
DT	01-MAR-2001 (TREMBLrel. 16, Last annotation update)	
DE	Alpha-1 Type XI COLLAGEN (FRAGMENT).	
OS	Gallus gallus (Chicken).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauvia; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.	
OC	NCBI_TAXID=9031;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STAIN=WHITE LEGHORN; TISSUE=LIMB;	
RX	MEDLINE=33054557; PubMed=1426667;	
Nah H.-D., Barenbaum M., Upholt W.B.;		
RA	361 NLVT 364	

ID	SEQUENCE	PRELIMINARY:	PRT:	1006 AA.	Matches	100:	Conservative	18:	Mismatches	60:	Indels	42:	Gaps	4:
Q9UUT5;					QY	41	PTENGLPGR--DGRDGREGPRGKGDGLPGMGLSGIQGPGP-----							82
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)				Db	1314	PGDGPPLGELGPGACQDGKEDGDCQGPGCSEAGPGPPKGRRPPGAAGAEGR							1373
DT	01-MAR-2001 (TREMBLrel. 16, Last annotation update)						COLLAGEN TYPE XI ALPHA-1 ISOFORM A.							
GN	COLLAGEN TYPE XI ALPHA-1 ISOFORM A.						COLLAGEN							
OS	Homo sapiens (Human).						COLLAGEN							
OC	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						COLLAGEN							
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.						COLLAGEN							
OX	NCBI-TAXID=9606;						COLLAGEN							
RN	[1]						COLLAGEN							
RP	SEQUENCE FROM N.A.						COLLAGEN							
RX	MEDLINE-20455728; PubMed-10486316;						COLLAGEN							
RA	Annen S., Korkko J., Czarny M., Warman M.L., Brunner H.G.,						COLLAGEN							
RA	Kaaijainen H., Mulliken J.B., Tranbäckberg L., Brooks D.G., Cox G.F.,						COLLAGEN							
RA	Crusberg J.R., Curtis M.A., Davenport S.L.H., Friedrich C.A.,						COLLAGEN							
RA	Kaitila I., Krawczynski M.R., Latoro-Bielenska A., Muital S.,						COLLAGEN							
RA	Olsen B.R., Shinno N., Somer M., Vilkula M., Zlotogora J.,						COLLAGEN							
RA	Prockop D.J., Ala-Kokko L.;						COLLAGEN							
RA	Kaitila I., Krawczynski M.R., Latoro-Bielenska A., Muital S.,						COLLAGEN							
RA	Olsen B.R., Shinno N., Somer M., Vilkula M., Zlotogora J.,						COLLAGEN							
RA	Prockop D.J., Ala-Kokko L.;						COLLAGEN							
RA	Kaitila I., Krawczynski M.R., Latoro-Bielenska A., Muital S.,						COLLAGEN							
RA	Olsen B.R., Shinno N., Somer M., Vilkula M., Zlotogora J.,						COLLAGEN							
RA	Prockop D.J., Ala-Kokko L.;						COLLAGEN							
RA	Kaitila I., Krawczynski M.R., Latoro-Bielenska A., Muital S.,						COLLAGEN							
RA	Olsen B.R., Shinno N., Somer M., Vilkula M., Zlotogora J.,						COLLAGEN							
RA	Prockop D.J., Ala-Kokko L.;						COLLAGEN							
RA	Kaitila I., Krawczynski M.R., Latoro-Bielenska A., Muital S.,						COLLAGEN							
RA	Olsen B.R., Shinno N., Somer M., Vilkula M., Zlotogora J.,						COLLAGEN							
RA	Prockop D.J., Ala-Kokko L.;						COLLAGEN							
RA	Kaitila I., Krawczynski M.R., Latoro-Bielenska A., Muital S.,						COLLAGEN							
RA	Olsen B.R., Shinno N., Somer M., Vilkula M., Zlotogora J.,						COLLAGEN							
RA	Prockop D.J., Ala-Kokko L.;						COLLAGEN							
RA	Kaitila I., Krawczynski M.R., Latoro-Bielenska A., Muital S.,						COLLAGEN							
RA	Olsen B.R., Shinno N., Somer M., Vilkula M., Zlotogora J.,						COLLAGEN							
RA	Prockop D.J., Ala-Kokko L.;						COLLAGEN							
RA	Kaitila I., Krawczynski M.R., Latoro-Bielenska A., Muital S.,						COLLAGEN							
RA	Olsen B.R., Shinno N., Somer M., Vilkula M., Zlotogora J.,						COLLAGEN							
RA	Prockop D.J., Ala-Kokko L.;						COLLAGEN							
RA	Kaitila I., Krawczynski M.R., Latoro-Bielenska A., Muital S.,						COLLAGEN							
RA	Olsen B.R., Shinno N., Somer M., Vilkula M., Zlotogora J.,						COLLAGEN							
RA	Prockop D.J., Ala-Kokko L.;						COLLAGEN							
RA	Kaitila I., Krawczynski M.R., Latoro-Bielenska A., Muital S.,						COLLAGEN							
RA	Olsen B.R., Shinno N., Somer M., Vilkula M., Zlotogora J.,						COLLAGEN							
RA	Prockop D.J., Ala-Kokko L.;						COLLAGEN							
RA	Kaitila I., Krawczynski M.R., Latoro-Bielenska A., Muital S.,						COLLAGEN							
RA	Olsen B.R., Shinno N., Somer M., Vilkula M., Zlotogora J.,						COLLAGEN							
RA	Prockop D.J., Ala-Kokko L.;						COLLAGEN							
RA	Kaitila I., Krawczynski M.R., Latoro-Bielenska A., Muital S.,						COLLAGEN							
RA	Olsen B.R., Shinno N., Somer M., Vilkula M., Zlotogora J.,						COLLAGEN							
RA	Prockop D.J., Ala-Kokko L.;						COLLAGEN							
RA	Kaitila I., Krawczynski M.R., Latoro-Bielenska A., Muital S.,						COLLAGEN							
RA	Olsen B.R., Shinno N., Somer M., Vilkula M., Zlotogora J.,						COLLAGEN							
RA	Prockop D.J., Ala-Kokko L.;						COLLAGEN							
RA	Kaitila I., Krawczynski M.R., Latoro-Bielenska A., Muital S.,						COLLAGEN							
RA	Olsen B.R., Shinno N., Somer M., Vilkula M., Zlotogora J.,						COLLAGEN							
RA	Prockop D.J., Ala-Kokko L.;						COLLAGEN							
RA	Kaitila I., Krawczynski M.R., Latoro-Bielenska A., Muital S.,						COLLAGEN							
RA	Olsen B.R., Shinno N., Somer M., Vilkula M., Zlotogora J.,						COLLAGEN							
RA	Prockop D.J., Ala-Kokko L.;						COLLAGEN							
RA	Kaitila I., Krawczynski M.R., Latoro-Bielenska A., Muital S.,						COLLAGEN							
RA	Olsen B.R., Shinno N., Somer M., Vilkula M., Zlotogora J.,						COLLAGEN							
RA	Prockop D.J., Ala-Kokko L.;						COLLAGEN							
RA	Kaitila I., Krawczynski M.R., Latoro-Bielenska A., Muital S.,						COLLAGEN							
RA	Olsen B.R., Shinno N., Somer M., Vilkula M., Zlotogora J.,						COLLAGEN							
RA	Prockop D.J., Ala-Kokko L.;						COLLAGEN							
RA	Kaitila I., Krawczynski M.R., Latoro-Bielenska A., Muital S.,						COLLAGEN							
RA	Olsen B.R., Shinno N., Somer M., Vilkula M., Zlotogora J.,						COLLAGEN							
RA	Prockop D.J., Ala-Kokko L.;						COLLAGEN							
RA	Kaitila I., Krawczynski M.R., Latoro-Bielenska A., Muital S.,						COLLAGEN							
RA	Olsen B.R., Shinno N., Somer M., Vilkula M., Zlotogora J.,						COLLAGEN							
RA	Prockop D.J., Ala-Kokko L.;						COLLAGEN							
RA	Kaitila I., Krawczynski M.R., Latoro-Bielenska A., Muital S.,						COLLAGEN							
RA	Olsen B.R., Shinno N., Somer M., Vilkula M., Zlotogora J.,						COLLAGEN							
RA	Prockop D.J., Ala-Kokko L.;						COLLAGEN							
RA	Kaitila I., Krawczynski M.R., Latoro-Bielenska A., Muital S.,						COLLAGEN							
RA	Olsen B.R., Shinno N., Somer M., Vilkula M., Zlotogora J.,						COLLAGEN							
RA	Prockop D.J., Ala-Kokko L.;						COLLAGEN							
RA	Kaitila I., Krawczynski M.R., Latoro-Bielenska A., Muital S.,						COLLAGEN							
RA	Olsen B.R., Shinno N., Somer M., Vilkula M., Zlotogora J.,						COLLAGEN							
RA	Prockop D.J., Ala-Kokko L.;						COLLAGEN							
RA	Kaitila I., Krawczynski M.R., Latoro-Bielenska A., Muital S.,						COLLAGEN							
RA	Olsen B.R., Shinno N., Somer M., Vilkula M., Zlotogora J.,						COLLAGEN							
RA	Prockop D.J., Ala-Kokko L.;						COLLAGEN							
RA	Kaitila I., Krawczynski M.R., Latoro-Bielenska A., Muital S.,						COLLAGEN							
RA	Olsen B.R., Shinno N., Somer M., Vilkula M., Zlotogora J.,						COLLAGEN							
RA	Prockop D.J., Ala-Kokko L.;						COLLAGEN							
RA	Kaitila I., Krawczynski M.R., Latoro-Bielenska A., Muital S.,						COLLAGEN							
RA	Olsen B.R., Shinno N., Somer M., Vilkula M., Zlotogora J.,						COLLAGEN							
RA	Prockop D.J., Ala-Kokko L.;						COLLAGEN							
RA	Kaitila I., Krawczynski M.R., Latoro-Bielenska A., Muital S.,						COLLAGEN							
RA	Olsen B.R., Shinno N., Somer M., Vilkula M., Zlotogora J.,						COLLAGEN							
RA	Prockop D.J., Ala-Kokko L.;						COLLAGEN							
RA	Kaitila I., Krawczynski M.R., Latoro-Bielenska A., Muital S.,						COLLAGEN							
RA	Olsen B.R., Shinno N., Somer M., Vilkula M., Zlotogora J.,						COLLAGEN							
RA	Prockop D.J., Ala-Kokko L.;						COLLAGEN							
RA	Kaitila I., Krawczynski M.R., Latoro-Bielenska A., Muital S.,						COLLAGEN							
RA	Olsen B.R., Shinno N., Somer M., Vilkula M., Zlotogora J.,						COLLAGEN							
RA	Prockop D.J., Ala-Kokko L.;						COLLAGEN							
RA	Kaitila I., Krawczynski M.R., Latoro-Bielenska A., Muital S.,						COLLAGEN							
RA	Olsen B.R., Shinno N., Somer M., Vilkula M., Zlotogora J.,						COLLAGEN							
RA	Prockop D.J., Ala-Kokko L.;						COLLAGEN							
RA	Kaitila I., Krawczynski M.R., Latoro-Bielenska A., Muital S.,						COLLAGEN							
RA	Olsen B.R., Shinno N., Somer M., Vilkula M., Zlotogora J.,						COLLAGEN							
RA	Prockop D.J., Ala-Kokko L.;						COLLAGEN							
RA	Kaitila I., Krawczynski M.R., Latoro-Bielenska A., Muital S.,						COLLAGEN							
RA	Olsen B.R., Shinno N., Somer M., Vilkula M., Zlotogora J.,						COLLAGEN							
RA	Prockop D.J., Ala-Kokko L.;						COLLAGEN							
RA	Kaitila I., Krawczynski M.R., Latoro-Bielenska A., Muital S.,						COLLAGEN							
RA	Olsen B.R., Shinno N., Somer M., Vilkula M., Zlotogora J.,						COLLAGEN							
RA	Prockop D.J., Ala-Kokko L.;						COLLAGEN							
RA	Kaitila I., Krawczynski M.R., Latoro-Bielenska A., Muital S.,						COLLAGEN							
RA	Olsen B.R., Shinno N., Somer M., Vilkula M., Zlotogora J.,						COLLAGEN							
RA	Prockop D.J., Ala-Kokko L.;						COLLAGEN							
RA	Kaitila I., Krawczynski M.R., Latoro-Bielenska A., Muital S.,						COLLAGEN							
RA	Olsen B.R., Shinno N., Somer M., Vilkula M., Zlotogora J.,						COLLAGEN							
RA	Prockop D.J., Ala-Kokko L.;						COLLAGEN							
RA	Kaitila I., Krawczynski M.R., Latoro-Bielenska A., Muital S.,						COLLAGEN							
RA	Olsen B.R., Shinno N., Somer M., Vilkula M., Zlotogora J.,						COLLAGEN							
RA	Prockop D.J., Ala-Kokko L.;						COLLAGEN							

Query Match 19 3%; Score 479.5; DB 4; Length 1838; Best Local Similarity 44.1%; Pred No. 9.8e-29; Matches 105; Conservative 9; Mismatches 67; Indels 57; gaps 4; Qy 41 PTENGTPRGDGRGEGKPGKPGCPGPMGSGLQGPTGVGVPGKENGSSAGEPCKGE 100 Db 1323 PGPKGPPGDPGPKSPGPGVCPGPGPQGDPGKDPGKDDGCEPQGDPG 1382 Qy 101 RGLSPPGPGPAGK-----EGPSGKQGNIGPO--GKPGP----- 136 Db 1383 PGPSGPPGRRGPPCPAGPESROEGKAGEAGLEGGPKTPIGPOGAPKPGDPGLRT 1442 Qy 137 ----- -KEAGKGEVGAPGMQGSTAKGSGVPGKE 166 Db 1443 PGPVGEQGIPGSPCPDPGPPGPMPPGLPGLKQGDPGKDDGCEPQGDPG 1502 Qy 167 RGAPGVQGA--PGNAGAAGPAGPAGQAGCSRSPGKGDGDRGKIKBSSLG 221 Db 1503 RGUPGPPQSSPKGEQDITGPSPGPICPGPGPGLKQGDPGKDDGCEPQGDPG 1560

Query Match 19 3%; Score 480; DB 4; Length 1818; Best Local Similarity 45.5%; Pred No. 8.8e-29; Matches 100; Conservative 18; Mismatches 60; Indels 42; Gaps 4; Qy 83 ----- VGPVKGENSGASAGERCPKPSGRGLGSPGSPGCPGKPG 1385 Db 1386 QGEKGAKGEAGAEGPKTGPGVCPGPGQPAKKGPGPEGLRGTGPGVQGLPGAAQDGPPG 1445 Qy 125 QSGNIPQCKPGKRGKGEAQPKGEWGAPEMNGEKGAKSTGPGERACVQGAP--GNAGA 181 Db 1446 --MGPPLGKGDPGSKGKGPHLGLIGLGPGEQEGKGDRLPGLQGSPCAKGDDGI 1502 Qy 182 AGPAGPAGPQGAPGSRQPPGKGDGKGPGDGRGKIKBSSLG 221 Db 1503 RGPAGPLGPGCPFGPLGPQGPKGNKGSTGAGQKGDGSP 1542

RESULT 15

015094 ID 015094 PRELIMINARY; PRT; 1838 AA.

AC 015094; DT 01-NOV-1996 (T=EMBrel. 01, Created) DT 01-NOV-1996 (T=EMBrel. 01, Last sequence update) DT 01-MAR-2001 (T=EMBrel. 16, Last annotation update) DE PRO-ALPHA-1 TYPE V COLLAGEN, COL5A1.

GN Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

OX NCBI_TaxID=9606; RN [1]

SEQUENCE FROM N.A. MEDLINE-92105142; GREENSPAN D.S., Cheng W., Hoffman G.G.; "The Pro-alpha 1(V) collagen chain. Complete primary structure, distribution of expression, and comparison with the pro-alpha 1(XI) collagen chain.", J. Biol. Chem. 266:24727-24733(1991).

DR InterPro; IPR00087; DR InterPro; IPR00088; DR InterPro; IPR001791; DR InterPro; IPR003129; DR Pfam; PF01391; collagen; 18. DR Pfam; PF01410; COLF1; 1. DR Pfam; PF02210; TSPN; 1. DR ProDom; PDB002078; SMART; SMART_00038; COF1; 1. DR SEQUENCE; 1838 AA; 183611 MW; 5078307F6E00FOBA CRC64;